Energy and related graph invariants

by

Eric Ould Dadah Andriantiana

Dissertation presented for the degree of Doctor of Philosophy in Mathematics in the Faculty of Sciences at Stellenbosch University

Department of Mathematical Sciences (Mathematics Division)
University of Stellenbosch
Private Bag X1, Matieland 7602 , South Africa.

Promoter: Prof. Stephan G. Wagner

December 2013

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
E. O. D. Andriantiana

Date: 2013/07/01 ....

Abstract

Energy and related graph invariants

E. O. D. Andriantiana

Department of Mathematical Sciences (Mathematics Division)
University of Stellenbosch
Private Bag X1, Matieland 7602 , South Africa.

Dissertation: PhD (Math.)

December 2013

In this thesis, a number of different graph invariants are studied in various subclasses of acyclic graphs and unicyclic graphs. This includes in particular the sum of the absolute values of the eigenvalues, which is known as the energy of a graph. Other spectral based invariants that we consider are the sum of exponentials of the eigenvalues, which is known as the Estrada index, and the spectral moments of a graph. The list extends to the Merrifield-Simmons index, the Hosoya index, the number of subtrees of a given order, the number of walks having a specified length, and several other related parameters.

We first consider the energy of trees. The n-vertex trees with largest, second-, third- and fourth-largest energy have been determined earlier for any given n. Estimating upper and lower bounds for the energies of trees with three leaves, and using the well-known technique of "sliding along a path", we are able to extend this list considerably for sufficiently large n.

Next we study the energy and two closely related parameters, namely the Hosoya- and the Merrifield-Simmons index, for trees with given degree sequence. The main result in this section is the characterisation of trees with given degree sequence D that minimizes the energy.

The last section on graph energy is concerned with unicyclic graphs: in a series of papers, various researchers aimed at finding the unicyclic graph with given number of vertices and largest energy. Their results narrowed the list of possible candidates to cycles and so-called "tadpole graphs", which are obtained by joining a vertex of a cycle to one of the ends of a path. We carefully estimate the energies of these special unicyclic graphs, which enables us to prove two conjectures on the largest and second-largest energy of unicyclic graphs (due to Caporossi, Cvetković, Gutman and Hansen and Gutman, Furtula and Hua respectively). An additional result characterizing the non-bipartite unicyclic graphs with largest energy is also proved.

In the last part of this thesis we show that the "greedy tree" with degree sequence D, which is constructed from a given degree sequence by a simple greedy algorithm, maximizes the number of subtrees of any given order and the k-th spectral moment for any nonnegative integer k among all trees whose degree sequence is majorized by D. We obtain a number of corollaries from this fact, most notably a conjecture of Ilić and Stevanović on trees with given maximum degree, which in turn implies a conjecture of Gutman, Furtula, Marković and Glišić on the Estrada index of such trees.

Opsomming

Energie en verwante grafiek invariants

("Energy and related graph invariants")

E. O. D. Andriantiana

Departement Wiskundige Wetenskappe (Wiskunde Afdeling,)
Universiteit Stellenbosch
Privaatsak X1, Matieland, 7602, Suid-Afrika

Proefskrif: PhD (Math.)

Desember 2013

In hierdie tesis word 'n verskeidenheid van grafiekinvariantes in verskillende subklasse van asikliese grafieke en unisikliese grafieke bestudeer. Dit sluit in die besonder die som van die absolute waardes van die eiewaardes, wat as die energie van 'n grafiek bekend staan, in. Ander spektraal-gebaseerde invariantes wat ons beskou is die som van eksponensiale van die eiewaardes, wat as die Estrada indeks bekend staan, en die spektrale momente van 'n grafiek. Die lys bevat verder die Merrifield-Simmons indeks, die Hosoya indeks, die aantal deelbome van 'n gegewe orde, die aantal wandelings met 'n gespesifiseerde lengte, en ander verwante parameters.

Ons beskou eers die energie van bome. Die n-nodus bome met die grootste, tweede-, derde- en vierde-grootste energie is al vroeër bepaal vir enige gegewe n. Afskattings van bo- en ondergrense vir die energie van bome met drie blare, en die gebruik van die bekende tegniek van "langs 'n pad gly", stel ons in staat om hierdie lys aansienlik uit te brei vir voldoende groot n.

Daarna bestudeer ons die energie en twee nou verwante parameters, naamlik die Hosoya- en die Merrifield-Simmons indeks, vir bome met 'n gegewe graadry. Die hoofresultaat in hierdie hoofstuk is die karakterisering van bome met gegewe graadry D wat die energie minimeer.

In die laaste gedeelte wat oor die energie van grafieke handel beskou ons unisikliese grafieke: in 'n reeks van artikels het verskeie navosers daarop gemik om die unisikliese grafiek met 'n gegewe aan-

tal nodusse en grootste energie te bepaal. Hulle resultate het die lys van moontlike kandidate vernou tot siklusse en die sogenaamde "paddavis-grafieke", wat verkry word deur 'n nodus van 'n siklus by een van die eindpunte van 'n pad aan te sluit. Ons skat die energie van hierdie spesiale unisikliese grafieke noukeurig af, wat ons in staat stel om twee vermoedes oor die grootste en tweede-grootste energie van unisikliese grafieke (onderskeidelik deur Caporossi, Cvetković, Gutman en Hansen en Gutman, Furtula en Hua) te bewys. 'N addisionele resultaat wat die nie-bipartiete unisikliese grafieke met die grootste energie karakteriseer word ook bewys.

In die laaste deel van hierdie tesis toon ons aan dat die "gulsige boom" met graadry D, wat uit 'n gegewe graadry deur 'n eenvoudige gulsige algoritme gebou word, die getal deelbome van enige gegewe orde en die k-te spektrale moment vir enige nie-negatiewe heelgetal k maksimeer onder al die bome wie se graadry deur D gemajoreer word. Ons kry 'n aantal gevolgtrekkings van hierdie feit, in die besonder 'n vermoede deur Ilić en Stevanović oor bome met gegewe maksimumgraad, wat 'n vermoede deur Gutman, Furtula, Marković en Glišić oor die Estrada indeks van bome impliseer.

Acknowledgements

I am very grateful to my supervisor Prof Stephan Wagner, for his help all along my postgraduate studies. It was very kind of him to accept to translate the abstract of this thesis in Afrikaans. I sincerely appreciate his support that allowed me to attend conferences and workshops to learn advanced skills and share parts of our results. In particular, some of the results in this thesis were presented in the following conferences and/or workshops:

- 30/05/2013 - 02/06/2013: International Workshop of Spectral Graph and Hypergraph Theory, Fuzhou, China,
- 03/09/2012 - 14/09/2012: ICTP-IPM Workshop and Conference in Combinatorics and Graph Theory, Trieste, Italy,
- 27/02/2012 - 23/03/2012: Block Course: Extremal Combinatorics in Random Discrete Structures, Berlin, Germany,
- 26/06/2011 - 02/07/2011: Algebraic Graph Theory Summer School, Rogla, Slovenia,
- 19/06/2011 - 25/06/2011: 7th Slovenian International Conference on Graph Theory, Bled, Slovenia.

I wish to extend my appreciation to:

- The German Academic Exchange Service (DAAD), whose generous support covered most of the financial cost of my PhD study.
- The University of Stellenbosch for hosting my postgraduate study and providing financial supports in many ways such as merit bursaries.
- Lecturers and colleagues who provided precious advice along my studies.

...by God's grace I am what I am, ... *Cor. I* 15:10

Dedications

To my dad Rasolofomanana Flaubert

Contents

List of Figures

Part I

Preliminaries

Chapter 1

Introduction

Chemical graph theory is a branch of mathematical chemistry. It is based on the possibility of describing the structure of molecules in terms of graphs: one such way is to represent atoms as vertices and the chemical bonds as edges. The resulting graph is called a *molecular graph*. From around the middle of the twentieth century, various interesting relations between graph-theoretical properties of molecular graphs and the physico-chemical properties of the corresponding molecules has been reported, and the literature of this field of study is still growing fast. Both graph theory and chemistry benefit from the interaction: many graph invariants were first introduced because of their chemical meaning, but they ended up attracting the attention of mathematicians and got their own place in pure mathematics. In this thesis, we will study several such graph invariants, aiming to characterize extremal graphs in specific classes of graphs.

A considerable portion of this work was conducted in collaboration with colleagues in South Africa and overseas, specifically Ivan Gutman, Hua Wang, Boris Furtula and Milan Cvetić.

Let G be a simple graph with set of vertices $\{v_1, \dots, v_n\}$, and let A be the adjacency matrix of G, that is, A is a square matrix whose entry at the crossing of i^{th} row and j^{th} column is equal to the number of edges (0 or 1) between v_i and v_j. The eigenvalues $\lambda_1, \dots, \lambda_n$ of A are also called eigenvalues of the graph G. The main graph invariants that we consider are the following:

i) The energy of a graph is defined to be the sum of the absolute values of its eigenvalues. Using the above notation, the energy of G is $\mathrm{En}(G) = |\lambda_1| + \cdots + |\lambda_1|$. The number of independent vertex subsets is known as Merrifield-Simmons index and denoted by σ. If G is as defined above, then $\sigma(G)$ is the number of subsets of $\{v_1, \dots, v_n\}$ which do not contain adjacent vertices. Similarly, the Hosoya index Z is defined as the number of independent edge subsets. These three parameters are studied in Part II:

In Chapter 3 we briefly mention historical notes on the three parameters, point out some of their applications in chemistry which motivate their study, and we also review some well-known relations between them (especially En and Z), which explain why we group the three in one part of this manuscript.

The class of trees is the most thoroughly studied class of graphs when it comes to those three parameters. But before 2011, only the n-vertex trees up to fourth-largest energy were known [28, 54, 64]. In Chapter 4, we show how to extend this list for sufficiently large n to contain all the trees up to $(3n-84)^{\text{th}}$ (odd n) or $(3n-87)^{\text{th}}$ (even n) largest energy. Such a list ends at the first appearance of a tree with more than three leaves. A similar list for trees with large Hosoya index and small Merrifield-Simmons index had already been obtained earlier [82] by Wagner. Motivated by the chemical background, we also studied [33] the case where n is not too large: estimating how long a part of the list we provided would be valid if $n \leq 100$. Chapter 4 corresponds to the articles [3, 33] published in MATCH Communications in Mathematical and in Computer Chemistry.

The class of trees with fixed degree sequence is considered in Chapter 5. We provide a theorem that characterizes the unique tree with a given degree sequence that has largest energy and Hosoya index and minimum Merrifield-Simmons index. Combined with a key observation comparing En, Z and σ for trees with different degree sequence, the theorem has as immediate corollaries the description of extremal trees in subclasses of trees obtained by adding extra conditions such as fixing diameter, number of leaves, and maximum degree. An article [4] based on Chapter 5 is published in Discrete Applied Mathematics.

Chapter 6 ends Part II. It consists of a study of the energy of unicyclic graphs, a natural class of graphs to study from a chemist's point of view: molecules whose structure contains a cycle occurs frequently. Among other things, we prove two conjectures due to Caporossi, Cvetković, Gutman and Hansen [11, 38] and Gutman, Furtula and Hua [34] respectively; they describe the n-vertex unicyclic graphs with maximum energy and second-largest energy. As additional results, we also showed theorems that characterize bipartite (resp. non-bipartite) unicyclic trees with order n and large energy. These results are published [6] in the journal Linear Algebra and its Applications.

ii) The number of subtrees and the spectral moment are investigated in Part III. The k^{th} spectral moment is defined by $\mathrm{M}_k(G) = \lambda_1^k + \cdots + \lambda_n^k$. Similar techniques are used to the study of the two invariants, and the structures of the extremal trees obtained are identical. The motivation for studying these invariants range from pure mathematical importance to applications in bioinformatics [63] and chemistry [37]. Spectral moments also relate the Estrada index [21] to the number of walks, a key relation often used in the study of Estrada index.

The greedy tree (constructed by assigning the highest degree to the root, the second-, third-, ... highest degrees to the neighbors of the root, and so on) has been shown to be an extremal tree in various papers on different invariants [9, 10, 75, 84, 90, 92]. Formal definitions of the greedy tree and several types of trees related to it are provided in Chapter 7, together with preliminary observations to be used in the two last chapters.

We manage to show in Chapter 8 that the greedy tree with degree sequence D maximizes the k^{th} spectral moment for any non-negative integer k among all trees whose degree sequence is majorized by D. A conjecture of Ilić and Stevanović on trees with given maximum degree follows as an immediate corollary of the main theorems in the chapter, which in turn implies a conjecture of Gutman, Furtula, Marković and Glišić on the Estrada index of such trees. More corollaries are pointed out as well. An article version [7] of Chapter 8 has been submitted for publication.

Chapter 9 concludes this thesis. It corresponds to the paper [8] published in the Electronic Journal of Combinatorics, where we strengthen the main result of [84]: we show that for any non-negative integer k the maximum number of subtrees of order k in a tree having a given degree sequence is obtained for the corresponding greedy tree. An additional theorem comparing the number of k-vertex subtrees of greedy trees with different degree sequences leads to various corollaries on extremal trees in several classes of trees, e. g. with fixed diameter or fixed number of leaves. Further results describing extremal trees with respect to the number of k-vertex subtrees containing the root or containing the root and having fixed number of leaves are also provided.

Chapter 2

Terminologies and basic notions

For the sake of completeness, we provide in this chapter definitions of standard terminologies, and we also mention some well-known properties of the notions defined.

A (simple undirected) *graph* G is characterized by its set of *vertices* and its set of *edges* respectively, denoted by $V(G)$ and $E(G)$, where an element of $E(G)$ consists of a two-element subset of $V(G)$. We then write $G = (V(G), E(G))$. The two vertices in an edge are its two *ends*. If $V(G)$ is empty, then so is $E(G)$, and G is called the *empty graph*. For simplicity, we write an edge with ends u and v as uv or vu instead of $\{u, v\}$. The *order* and the *size* of G are respectively $|V(G)|$ and $|E(G)|$. In a graphical representation of a graph, each of its vertices is represented by a dot and each of its edges is represented by a continuous line, see Figure 2.2. A graph H is a *subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, we write $H \subseteq G$. If furthermore $H \neq G$, then H is a *proper subgraph* of G. A *path* is a graph of the form $(\{v_1, \dots, v_n\}, \{v_i v_{i+1} : i \in [1, \dots, n-1]\})$ where the v_is are distinct; a *cycle* is a graph of the form $(\{v_1, \dots, v_n\}, \{v_i v_{i+1 \mod n} : i \in [1, \dots, n]\})$ for distinct v_is (Figure 2.1). The *length* of a path or a cycle is its

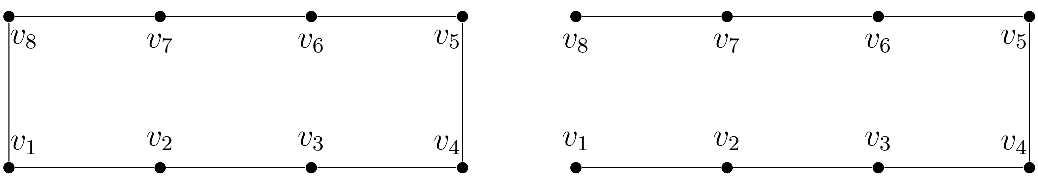


Figure 2.1: A cycle (left-hand side) and a path (right-hand side) with 8 vertices

number of edges. The length of the longest path in a graph G is called *diameter* of G, it is denoted by $\mathrm{diam}(G)$. G is *connected* if for any two elements u and v of $V(G)$ there is a path in G which starts from u and ends at v. An *acyclic graph*, also called a *forest*, is a graph which does not contain a cycle. A *tree* is a connected forest. In a tree, a vertex whose degree is 1 is called *leaf*. In some situation it happens to be convenient to consider one vertex or one edge of a tree as special; such a vertex or edge is then called the *root* of the tree. A tree for which a fixed root has been chosen is a *rooted tree*. A *rooted forest* is a forest whose components are rooted trees. The root of a rooted tree T is denoted by $\mathrm{r}(T)$. The *height* of a rooted forest F, denoted by $\mathrm{h}(F)$, is the length of a longest path of F starting from a vertex root or from an end vertex of an edge root but does not use the edge root.

For any subgraph H of G, we define $G - H$ to be the subgraph of G whose set of vertices and set of edges are $V(G) - V(H)$ and $E(G) - \{e \in E(G) : e \cap V(H) \neq \varnothing\}$, respectively. For any edge $e \in E(G)$ we have $G - e := (V(G), E(G) - \{e\})$.

We say that two vertices u and v are *adjacent* if uv is an edge. Two distinct edges uv and $u'v'$ are also said to be adjacent if the intersection of $\{u, v\}$ and $\{u', v'\}$ is not empty. The edge uv is *incident* to u and to v. The set of all vertices of G adjacent to a vertex v is denoted by $N_G(v)$ or simply $N(v)$ if it is clear from the context which graph we are working on, that is $N_G(v) := \{u \in V(G) : vu \in E(G)\}$. $N_G(v)$ is called the *neighborhood* of v in G. The number $|N_G(v)|$ of vertices of G adjacent to v is the *degree* of v, it is denoted by $\deg_G(v)$ or simply $\deg(v)$. If $d_1 \geq \cdots \geq d_n$ are the degrees of the vertices in an n-vertex graph G, then the n-tuple $(d_1, \ldots, d_n)$ is called the *degree sequence* of G. If $(d_1, \ldots, d_n)$ and $(b_1, \ldots, b_n)$ are two degree sequences of graphs, we say that $D = (d_1, \ldots, d_n)$ *majorizes* $B = (b_1, \ldots, b_n)$ if and only if for each $k \in \{1, \ldots, n\}$ we have

$$\sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} d_i,$$

and we write $B \preccurlyeq D$.

Two graphs G and G' are *isomorphic* if there exists a bijective map $f : V(G) \longrightarrow V(G')$ such that $uv \in E(G)$ if and only if $f(u)f(v) \in E(G')$. Most of the time we identify isomorphic graphs. For instance, whenever A is a set of cardinality n then we denote by K_n the graph $(A, \{\{u, v\} : \{u, v\} \subseteq A\})$ which has n vertices and all possible edges between them. Such a graph is called *complete graph*. Similarly we denote by $\overline{K}_n := (A, \varnothing)$ the n-vertex edgeless graph.

In order to benefit from the rich literature of linear algebra, it is very common in graph theory to encode graphs by matrices. Associated to any given n-vertex graph G with set of vertices $V(G) =$

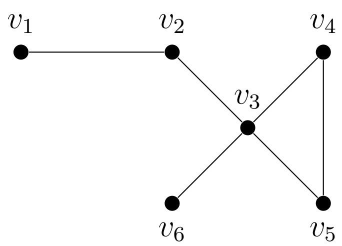
$\{v_1, \ldots, v_n\}$, we define an $n \times n$ matrix $A(G)$, where the entry $(A(G))_{i,j}$ of $A(G)$ at the crossing of the i^{th} row counted from top to bottom and the j^{th} column counted from left to right is given by

$$(A(G))_{i,j} = \begin{cases} 1 \text{ if } v_i v_j \in E(G), \\ 0 \text{ otherwise.} \end{cases}$$

See Figure 2.2 for an example. $A(G)$ is the *adjacency matrix* of G. Different adjacency matrices can possibly be obtained by different choices of labeling of the vertices of the concerned graph. But any two adjacency matrices, say A and B, of the same graph are identical up to permutation of rows and columns. More precisely, there exists a matrix P obtained by appropriately permuting rows in the $n \times n$ identity matrix I_n such that $A = PBP^{-1}$. Therefore they have the same characteristic polynomial

$$\begin{aligned} \det(xI_n - A) &= \det(xI_n - PBP^{-1}) \\ &= \det(xPI_nP^{-1} - PBP^{-1}) \\ &= \det P(xI_n - B)P^{-1} \\ &= \det(xI_n - B). \end{aligned}$$

Among other types of matrices associated to graphs, the adjacency



$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Figure 2.2: Graphical representation and adjacency matrix of the graph $G = (\{v_1, v_2, v_3, v_4, v_5, v_6\}, \{v_1v_2, v_2v_3, v_3v_4, v_4v_5, v_5v_3, v_3v_6\})$

matrix is so popular that the *characteristic polynomial* and the *eigenvalues* of a graph are defined to be those of its adjacency matrix. The characteristic polynomial of G is denoted by $\Phi(G, x)$.

We also define a few chemical terminologies that we will eventually use. In organic chemistry, a *hydrocarbon* is an organic compound consisting entirely of hydrogen and carbon. A hydrocarbon which contains two or more double bonds alternating with single bonds are called *conjugated hydrocarbons*.

For undefined notations and terminology, the reader is referred to [16].

Part II

Energy and independent subsets

Chapter 3

Introduction and basic notions

Let G be a graph. Two edges or two vertices of G which are not adjacent to each other are said to be *independent*. Any subset of $V(G)$ which does not contain two adjacent vertices is an *independent vertex subset* of G. The number of independent vertex subsets of G, also known as σ-index of G, is denoted by $\sigma(G)$. A subset of $E(G)$ is called an *independent edge subset* or *matching* of G if it does not contain adjacent edges. The number of independent edge subsets of cardinality k in G is denoted by $\mathrm{m}(G,k)$. The total number of matchings of G is denoted by $\mathrm{Z}(G)$, this means

$$\mathrm{Z}(G) = \sum_{k\geq 0} \mathrm{m}(G,k). \tag{3.1}$$

Z is also called the Z-index. Note that the empty set is an independent vertex subset and an independent edge subset of any graph. Hence $\sigma(G)$ and $\mathrm{Z}(G)$ can never be zero, no matter what G is.

The Z-index was introduced by the Japanese chemist Haruo Hosoya in a paper [44] published in 1971, where he studied among other things the relation between Z and characteristic polynomials of graphs representing carbon skeletons of saturated hydrocarbons, where each vertex represents a carbon atom and each edge represents a chemical bond. We call such a graph molecular graph. Hosoya also pointed out in his paper a correlation between the Z-index of a graph and the boiling point of the corresponding molecule. Further research was done later strengthening his observations, it was shown that in this context of graphs representing carbon skeletons of molecules the Z-index of a graph helps predicting more physicochemical properties of the corresponding molecule such as octane number [45], π-electron system [39, 40, 48] and others [27, 49, 72].

The history of the σ-index can be traced back to a paper [71] written by Richard E. Merrifield and Howard E. Simmons, which appeared in 1980. They present in the same paper the correlation between the σ-index of molecular graphs and the boiling point of the corresponding molecules. Both Z and σ have rich connections to well-known mathematical notions, namely Fibonacci numbers, Lucas numbers and Pascal's triangle [46, 47, 74]. For example, if $(F_i)_{i\in\mathbb{N}}$ is the sequence of the Fibonacci numbers, where $F_1 = F_2 = 1$ and $F_i+F_{i+1} = F_{i+2}$ for all $i \in \mathbb{N}$, then we have $\sigma(P_n) = F_{n+2}$ and $\mathrm{Z}(P_n) = F_{n+1}$ for all $n \in \mathbb{N}$. This is why σ is also called *Fibonacci number of a graph*. Nowadays, Z and σ are usually called *Hosoya index* and *Merrifield-Simmons index* respectively to honor their inventors.

The following elementary properties of Z and σ express the numbers of independent subsets of a graph in terms of those of smaller graphs. They play important roles in the rest of this thesis, where reasoning by induction is often used.

Lemma 3.1 *Let G and G' be two disjoint graphs. Then we have*

$$\mathrm{Z}(G \cup G') = \mathrm{Z}(G)\,\mathrm{Z}(G'), \tag{3.2}$$

$$\sigma(G \cup G') = \sigma(G)\sigma(G'). \tag{3.3}$$

If $v \in V(G)$, then we have

$$\mathrm{Z}(G) = \mathrm{Z}(G - v) + \sum_{w\in N_G(v)} \mathrm{Z}(G - \{v, w\}), \tag{3.4}$$

$$\sigma(G) = \sigma(G - v) + \sigma(G - (\{v\} \cup N_G(v))). \tag{3.5}$$

Proof. Choosing an independent edge subset S of $G \cup G'$ amounts to choosing an independent edge subset $S \cap V(G)$ from G and an independent edge subset $S \cap V(G')$ from G'. Thus we have $\mathrm{Z}(G \cup G') = \mathrm{Z}(G)\,\mathrm{Z}(G')$. The same idea applied to σ leads to equation (3.3).

(3.4) and (3.5) first count independent edge or vertex subsets which do not cover v and then count the rest separately. $\square$

The idea of the above proof also leads to

$$\mathrm{m}(G \cup G', k) = \sum_{i+j=k} \mathrm{m}(G, i)\,\mathrm{m}(G', j) \tag{3.6}$$

and

$$\mathrm{m}(G, k) = \mathrm{m}(G - v, k) + \sum_{w\in N_G(v)} \mathrm{m}(G - \{v, w\}, k - 1) \tag{3.7}$$

for any $k \in \mathbb{N}$. Given an arbitrary positive integer k, by counting the number of matchings of order k covering uv first and then counting the other matchings of order k we obtain:

$$\mathrm{m}(G, k) = \mathrm{m}(G - \{u, v\}, k - 1) + \mathrm{m}(G - uv, k). \tag{3.8}$$

Definition 3.2 Let G be a graph of order n and let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues. The graph invariant

$$\mathrm{En}(G) = \sum_{k=1}^{n} |\lambda_k| \tag{3.9}$$

is called the *energy* of G.

Note that the adjacency matrix of any graph is symmetric and has zero diagonal, hence the eigenvalues of a graph are always real and sum up to zero. Therefore (3.9) means that

$$\mathrm{En}(G) = 2 \sum_{\substack{\lambda \in \{\lambda_1, \ldots, \lambda_n\} \\ \lambda > 0}} \lambda = -2 \sum_{\substack{\lambda \in \{\lambda_1, \ldots, \lambda_n\} \\ \lambda < 0}} \lambda.$$

Studies involving the sum of the negative eigenvalues of graphs can already be found in the literature in the 1940's. For example, the British applied mathematician and theoretical chemist Charles A. Coulson established [14] an interesting relation between the the sum of the negative eigenvalues of molecular graphs representing conjugated hydrocarbons and the total energy of the mobile electrons, also called π-electron energy, in the molecules. Most of the subsequent results following this line of research appeared in the 1970's or later, see for instance [37, 39, 48]. An approximation of the total π-electron energy $\mathcal{E}_\pi$ of a conjugated hydrocarbon molecule, whose molecular graph is G, using the method of Hückel molecular orbital [13] is given by the formula [31]

$$\mathcal{E}_\pi = n\alpha + \beta\,\mathrm{En}(G),$$

for some constants α and β, where n is the number of carbon atoms in the molecule. The discovery of this formula added more motivation to the study of En which went far beyond the class of molecular graphs. It later received its own name as *energy of a graph*, a name suggested [29] by Gutman in 1978. Pure mathematicians joined chemists in studying the energy of graphs in several classes of graphs which may or may not include non-molecular graphs, see [4, 6, 11, 53, 86]. We refer the reader to the text book [68] for a wide range of references on the energy of graphs.

A breakthrough was made by Coulson [14] when he showed how complex analysis can be used as a tool in the study of the energy of graphs. He provided an alternative integral formula for En:

Theorem 3.3 ([14]) *Let $\Phi'(G,x)$ be the first derivative of the characteristic polynomial $\Phi(G,x)$ of G. Then we have*

$$\begin{aligned}\mathrm{En}(G) &= \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[n - \frac{\mathbf{i}x\Phi'(G,\mathbf{i}x)}{\Phi(G,\mathbf{i}x)}\right]dx \\ &= \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[n - x\frac{d}{dx}\log\Phi(G,\mathbf{i}x)\right]dx, \end{aligned} \tag{3.10}$$

where $\mathrm{i} = \sqrt{-1}$.

Several varieties of integral formulas for the energy were found later as corollaries of Theorem 3.3.

Corollary 3.4 ([31]) *If G_1 and G_2 are two graphs with equal number of vertices, then*

$$\mathrm{En}(G_1) - \mathrm{En}(G_2) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\log\frac{\Phi(G_1,\mathbf{i}x)}{\Phi(G_2,\mathbf{i}x)}dx. \tag{3.11}$$

Because $\mathrm{En}(G_1) - \mathrm{En}(G_2)$ is a real number, it must be equal to the real part of the right hand side of (3.11). This means

$$\mathrm{En}(G_1) - \mathrm{En}(G_2) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\log\left|\frac{\Phi(G_1,\mathbf{i}x)}{\Phi(G_2,\mathbf{i}x)}\right|dx. \tag{3.12}$$

Let G be an n-vertex graph, and write its characteristic polynomial as

$$\Phi(G,x) = \sum_{k=0}^{n} a_k x^{n-k}.$$

Since $\Phi(\overline{K_n}) = x^n$ and $\mathrm{En}(\overline{K_n}) = 0$ for any positive integer n, we can deduce by using (3.12)

$$\begin{aligned}
\mathrm{En}(G) = \mathrm{En}(G) - \mathrm{En}(\overline{K_n}) &= \frac{1}{\pi}\int_{-\infty}^{+\infty} \log\left|\frac{\sum_{k=0}^{n} a_k(\mathbf{i}x)^{n-k}}{(\mathbf{i}x)^n}\right| dx \\
&= \frac{1}{\pi}\int_{-\infty}^{+\infty} \log\left|\sum_{k=0}^{n} a_k(\mathbf{i}x)^{-k}\right| dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \log\left[\left(\sum_{k\geq 0}(-1)^k a_{2k}x^{-2k}\right)^2 + \left(\sum_{k\geq 0}(-1)^k a_{2k+1}x^{-(2k+1)}\right)^2\right] dx \\
&= \frac{1}{\pi}\int_{0}^{+\infty} \log\left[\left(\sum_{k\geq 0}(-1)^k a_{2k}x^{-2k}\right)^2 + \left(\sum_{k\geq 0}(-1)^k a_{2k+1}x^{-(2k+1)}\right)^2\right] dx.
\end{aligned}$$

After a change of variable $z = 1/x$ we have the well-known relation [68]

$$\begin{aligned}
&\mathrm{En}(G) \\
&= \mathrm{En}(G) - \mathrm{En}(\overline{K_n}) \\
&= \frac{1}{\pi}\int_{0}^{+\infty} \frac{1}{z^2}\log\left[\left(\sum_{k\geq 0}(-1)^k a_{2k}z^{2k}\right)^2 + \left(\sum_{k\geq 0}(-1)^k a_{2k+1}z^{2k+1}\right)^2\right] dz. \qquad (3.13)
\end{aligned}$$

In a book [36] by Gutman and Polanski, a version of (3.13) specifically for trees is obtained. This formula also shows how the energy is related to the number of matchings. It reads as follows:

Theorem 3.5 ([36]) *If T is a tree with n vertices, then*

$$\Phi(T, x) = \sum_{k\geq 0}(-1)^k \,\mathrm{m}(T, k)x^{n-2k}$$

and hence

$$\mathrm{En}(T) = \frac{2}{\pi}\int_0^{\infty} \frac{1}{x^2}\log\left(\sum_{k\geq 0}\mathrm{m}(T, k)x^{2k}\right) dx. \qquad (3.14)$$

The formula (3.14) suggests to consider another graph invariant defined by

$$\mathrm{M}(T, x) := \sum_{k\geq 0}\mathrm{m}(T, k)x^k$$

for any tree T, so that we have

$$\mathrm{Z}(T) = \mathrm{M}(T, 1) \qquad (3.15)$$

and

$$\mathrm{En}(T) = \frac{2}{\pi}\int_0^\infty \frac{1}{x^2}\log \mathrm{M}(T,x^2)dx. \tag{3.16}$$

$\mathrm{M}(T,x)$ is the *matching polynomial* of T.

Remark 3.6 The importance of M is based on the fact that if T and T' are trees and $\mathrm{M}(T,x) \le \mathrm{M}(T',x)$ for all positive real numbers x (in particular $\mathrm{M}(T,x^2) \le \mathrm{M}(T',x^2)$ for all real x), then $\mathrm{En}(T) \le \mathrm{En}(T')$ and $\mathrm{Z}(T) \le \mathrm{Z}(T')$. If furthermore, there exists a real number $x > 0$ such that $\mathrm{M}(T,x) < \mathrm{M}(T',x)$, then we have $\mathrm{En}(T) < \mathrm{En}(T')$.

We have a more general version of Lemma 3.1 in terms of M.

Lemma 3.7 *Let G and G' be two disjoint graphs and let $x > 0$ be a real number. Then we have*

$$\mathrm{M}(G \cup G', x) = \mathrm{M}(G,x)\,\mathrm{M}(G',x). \tag{3.17}$$

If $v \in V(G)$, then we have

$$\mathrm{M}(G,x) = \mathrm{M}(G-v,x) + x\sum_{w\in N_G(v)} \mathrm{M}(G-\{v,w\},x). \tag{3.18}$$

For any $uv \in E(G)$ we have

$$\mathrm{M}(T,x) = \mathrm{M}(T-uv,x) + x\,\mathrm{M}(T-\{u,v\},x). \tag{3.19}$$

Proof. Let G, G', v and x be as described in the statement of the lemma. Using (3.6) we obtain

$$\begin{aligned}
\mathrm{M}(G\cup G',x) &= \sum_{k\ge 0} \mathrm{m}(G\cup G',k)x^k \\
&= \sum_{k\ge 0}\sum_{\substack{i+j=k\\ i,j\ge 0}} \mathrm{m}(G,i)\,\mathrm{m}(G',j)x^{i+j} \\
&= \sum_{i\ge 0}\mathrm{m}(G,i)x^i \sum_{j\ge 0}\mathrm{m}(G',i)x^j \\
&= \mathrm{M}(G,x)\,\mathrm{M}(G',x).
\end{aligned}$$

Use of equation (3.7) leads to

$$\begin{aligned}
\mathrm{M}(G,x) &= \sum_{k\ge 0}\mathrm{m}(G,k)x^k \\
&= \sum_{k\ge 0}\mathrm{m}(G-v,k)x^k + \sum_{k\ge 0}\sum_{w\in N_G(v)} \mathrm{m}(G-\{v,w\},k-1)x^k \\
&= \mathrm{M}(G-v,x) + x\sum_{w\in N_G(v)}\mathrm{M}(G-\{v,w\},x).
\end{aligned}$$

Using (3.8) we obtain

$$\begin{aligned}
\mathrm{M}(G,x) &= \sum_{k\geq 0} \mathrm{m}(G,k)x^k \\
&= \sum_{k\geq 0} (\mathrm{m}(G-\{u,v\},k-1) + \mathrm{m}(G-uv,k))x^k \\
&= x\sum_{k\geq 0} \mathrm{m}(G-\{u,v\},k)x^k + \sum_{k\geq 0} \mathrm{m}(G-uv,k)x^k \\
&= x\,\mathrm{M}(G-\{u,v\},x) + \mathrm{M}(G-uv,x).
\end{aligned}$$

□

A similar lemma is also known for the characteristic polynomial of graphs:

Lemma 3.8 ([15]) *Let uv be an edge of a graph G. Then*

$$\Phi(G,x) = \Phi(G-uv,x) - \Phi(G-\{u,v\},x) - 2\sum_{C\in\mathcal{C}(uv)} \Phi(G-C,x),$$

where $\mathcal{C}(uv)$ is the set of cycles containing uv. In particular if uv is a pendant edge with pendent vertex v, then

$$\Phi(G,x) = x\Phi(G-v,x) - \Phi(G-\{u,v\},x).$$

It is convenient to set $\Phi(\varnothing,x)=1$. Then both formulas in Lemma 3.8 remain correct.

Chapter 4

Trees with Large Energy

Let $T(i, j, n-i-j-1)$ denote the n-vertex tripod which has two branches of length i and j, respectively (see Figure 4.1). One can rearrange i, j and $n-i-j-1$ if needed and still have the same tripod. We call the only vertex of degree 3 in $T(i, j, n-i-j-1)$ the *center*. The four n-vertex trees with maximum energy, for $n \geq$

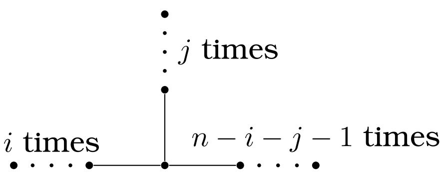


Figure 4.1: Tripod $T(i, j, n-i-j-1)$

15, are $P_n, T(2, 2, n-5), T(2, 4, n-7)$ and $T(2, 6, n-9)$, in this order [28, 54, 64]. The main purpose of this chapter is to extend this list, for large enough n, until the first appearance of a tree with four leaves, which is the tree with $(3n-84)^{\text{th}}$ (resp. $(3n-87)^{\text{th}}$) largest energy for odd n (resp. even n). Similar results for the parameters σ and Z were obtained by Wagner [82]: he determined the lists of trees with large Hosoya index and small Merrifield-Simmons index, respectively, from the path until the first appearance of a tree with four leaves. To achieve our goal, we use a similar approach as in [41, 82, 95]: the technique consists of considering a graph obtained by attaching a subgraph G to the i^{th} vertex in a path, and then observing how the energy depends on the choice of the position i. This technique will be briefly reviewed in Section 4.1. In Section 4.2 we show for all $n \geq 10$ that the tree $H(2, 2, 2, 2, n)$, obtained by merging each end of P_{n-8} to the third vertex in a 5-vertex path, is the tree with maximum energy among all trees of order n with at least four leaves. Section 4.3 is devoted to the comparison of the energy of

$H(2, 2, 2, 2, n)$ with those of tripods. The main theorem of this chapter is provided in Section 4.4, which consists of an ordering of the tripods with larger energy than $H(2, 2, 2, 2, n)$. Since the theorem only holds for large enough n, we also check the case of small values of n ranging from 10 to 100, and compare the lists of the corresponding trees with largest energy with the list in the main theorem.

4.1 "Sliding along a path"

Let G be a connected graph with at least two vertices, and let v be a vertex of G. Let n and k be integers such that $n-1 \geq k \geq 0$. We denote by $P(n, k, G, v)$ the graph which results from identifying v with the $(k+1)^{\text{th}}$ vertex in an n-vertex path, as in Figure 4.2. We aim to

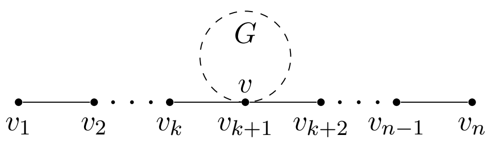


Figure 4.2: $P(n, k, G, v)$

understand how $\mathrm{M}(P(n, k, G, v), x)$ behaves as a function of k.

The following ordering of the $P(n, k, G, v)$s is well-known, see [41] and [82].

Lemma 4.1 *Let x be a positive real number and $n \geq 7$ an integer. Then the following inequalities hold:*

$$\begin{aligned}
&\mathrm{M}(P(n,0,G,v),x) > \mathrm{M}(P(n,2,G,v),x) > \cdots \\
&\cdots > \mathrm{M}(P(n,2\lfloor (n-1)/4 \rfloor,G,v),x) > \mathrm{M}(P(n,2\lfloor (n+1)/4 \rfloor -1,G,v),x) > \cdots \\
&\qquad\qquad\qquad \cdots > \mathrm{M}(P(n,3,G,v),x) > \mathrm{M}(P(n,1,G,v),x).
\end{aligned}$$

The values of k in the $P(n, k, G, v)$s increase from 0 to $2\lfloor (n-1)/4 \rfloor$ and then decrease from $2\lfloor (n+1)/4 \rfloor - 1$ to 1, by steps of 2.

An alternative proof for this is also presented in the master's thesis [1]. Note that $2\lfloor m/4 \rfloor$ and $2\lfloor (m+2)/4 \rfloor - 1$ are the two largest integers less or equal to $m/2$ for all positive integers m. This is because if $m \equiv 0, 1 \mod 4$ then

$$\frac{m}{2} - 2 = 2\left(\frac{m+2}{4} - 1\right) - 1 < 2\left\lfloor \frac{m+2}{4} \right\rfloor - 1 = 2\left\lfloor \frac{m}{4} \right\rfloor - 1 < 2\left\lfloor \frac{m}{4} \right\rfloor \leq \frac{m}{2},$$

and for $m \equiv 2, 3 \mod 4$ we have

$$\frac{m}{2} - 2 = 2\left(\frac{m}{4} - 1\right) < 2\left\lfloor\frac{m}{4}\right\rfloor < 2\left(\left\lfloor\frac{m}{4}\right\rfloor + 1\right) - 1$$
$$= 2\left\lfloor\frac{m+2}{4}\right\rfloor - 1 \leq 2\frac{m+2}{4} - 1 = \frac{m}{2}.$$

As k varies, G appears to be "sliding" along the path to which it is attached. This is the reason why lemmas of such a type are also called "Sliding along a path" [83].

The following remark is an immediate consequence of Lemma 4.1, it is particularly useful in practice to construct trees with larger energy than a given one (see [30] for instance).

Remark 4.2 The graph transformation described in Figure 4.3 reduces the number of leaves and increases the energy, for all integers $n > k > 1$. In general, the energy of a tree increases if we replace a

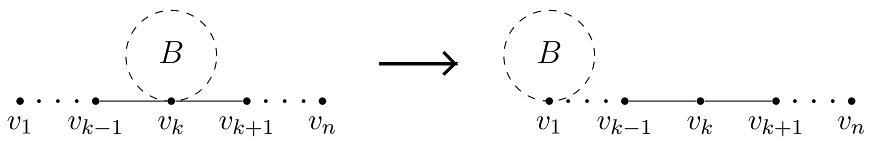


Figure 4.3: Moving B from v_k to v_1

branch which is not a path by a path of the same order.

By considering one of the branches of a tripod as a sliding subgraph, the following theorem follows from Remark 3.6 and Lemma 4.1:

Theorem 4.3 ([28]) *For all positive integers i and n such that $n \geq 3i + 7$ we have*

$$\mathrm{En}(T(i, 2\lceil i/2\rceil, n-i-2\lceil i/2\rceil - 1)) > \mathrm{En}(T(i, 2\lceil i/2\rceil + 2, n-i-2\lceil i/2\rceil - 3))$$
$$> \cdots > \mathrm{En}(T(i, 2\lfloor (n-i-1)/4\rfloor, n-i-2\lfloor (n-i-1)/4\rfloor - 1))$$
$$> \mathrm{En}(T(i, 2\lfloor (n-i+1)/4\rfloor - 1, n-i-2\lfloor (n-i+1)/4\rfloor)) > \cdots >$$
$$\mathrm{En}(T(i, 2\lfloor i/2\rfloor + 3, n-i-2\lfloor i/2\rfloor - 4)) > \mathrm{En}(T(i, 2\lfloor i/2\rfloor + 1, n-i-2\lfloor i/2\rfloor - 2)).$$

In the first two lines the length of the second shortest branch increases at each step by 2 until it reaches $2\lfloor (n-i-1)/4\rfloor$, and in the two last lines it decreases from $2\lfloor (n-i+1)/4\rfloor - 1$ to $2\lfloor i/2\rfloor + 1$.

4.2 Trees with at least four leaves and maximum energy

Throughout this section $d_{11}, d_{12}, d_{21}, d_{22}$ are always positive integers. For all integers n such that $d_{11} + d_{12} + d_{21} + d_{22} \leq n - 1$, we denote by $H(d_{11}, d_{12}, d_{21}, d_{22}, n)$ the n-vertex quadripod as described in Figure 4.4. It is convenient to set

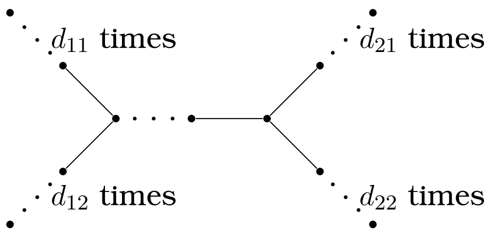


Figure 4.4: $H(d_{11}, d_{12}, d_{21}, d_{22}, n)$

$$T(0, j, k) = P_{j+k+1},$$
$$T(-1, j, k) = P_j \cup P_k$$

and

$$\mathrm{M}(T(-2, j, k), x) = \mathrm{M}(P_j, x)\,\mathrm{M}(P_{k-1}, x) + \mathrm{M}(P_{j-1}, x)\,\mathrm{M}(P_k, x)$$

for all positive integers j and k so that the well-known relations

$$\mathrm{M}(T(i, j, k+2), x) = \mathrm{M}(T(i, j, k+1), x) + x\,\mathrm{M}(T(i, j, k), x) \tag{4.1}$$

and

$$\begin{aligned}
&\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}, n), x)\\
&= \mathrm{M}(P_{d_{11}}, x)\,\mathrm{M}(P_{d_{12}}, x)\,\mathrm{M}(T(d_{21}, d_{22}, n - d_{11} - d_{12} - d_{21} - d_{22} - 2), x)\\
&+ x\,\mathrm{M}(P_{d_{11}-1}, x)\,\mathrm{M}(P_{d_{12}}, x)\,\mathrm{M}(T(d_{21}, d_{22}, n - d_{11} - d_{12} - d_{21} - d_{22} - 2), x)\\
&+ x\,\mathrm{M}(P_{d_{11}}, x)\,\mathrm{M}(P_{d_{12}-1}, x)\,\mathrm{M}(T(d_{21}, d_{22}, n - d_{11} - d_{12} - d_{21} - d_{22} - 2), x)\\
&+ x\,\mathrm{M}(P_{d_{11}}, x)\,\mathrm{M}(P_{d_{12}}, x)\,\mathrm{M}(T(d_{21}, d_{22}, n - d_{11} - d_{12} - d_{21} - d_{22} - 3), x),
\end{aligned} \tag{4.2}$$

which are obtained by applying (3.18), are valid for $i \geq -2$, $j, k \geq 1$ and $n \geq d_{11} + d_{12} + d_{21} + d_{22} + 1$. Equation (4.2) shows (using (4.1)) that

$$\begin{aligned}
&\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}, n+2), x)\\
&= \mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}, n+1), x) + x\,\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}, n), x)
\end{aligned} \tag{4.3}$$

and

$$\begin{aligned}
&\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22} + 2, n+2), x)\\
&= \mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22} + 1, n+1), x) + x\,\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}, n), x).
\end{aligned} \tag{4.4}$$

Lemma 4.4 *The n-vertex tree with at least four leaves and maximum energy must be an element of*

$$\{H(1,1,1,1,n), H(1,1,2,d_{22},n), H(2,d_{12},2,d_{22},n) : d_{22} \geq d_{12} \geq 1\}.$$

Proof. Remark 4.2 reduces the set of candidates to be the set of quadripods. Using the Lemma of "Sliding along a path" we know that if $\max\{d_{11}, d_{12}\} \geq 3$ and $\min\{d_{11}, d_{12}\} \neq 2$, then for all positive x we have

$$\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}), x) < \mathrm{M}(H(2, d_{12} + d_{11} - 2, d_{21}, d_{22}), x).$$

Similarly, if $\max\{d_{21}, d_{22}\} \geq 3$ and $\min\{d_{21}, d_{22}\} \neq 2$, then we have

$$\mathrm{M}(H(d_{11}, d_{12}, d_{21}, d_{22}), x) < \mathrm{M}(H(d_{11}, d_{12}, 2, d_{22} + d_{21} - 2), x).$$

Therefore, we must have $2 \in \{d_{i1}, d_{i2}\}$ or $d_{i1} = d_{i2} = 1$ for $i \in \{1, 2\}$. □

Now we also use (3.18) to get the following relations:

$$\begin{aligned}\mathrm{M}(H(1{,}1{,}1{,}1{,}n), x) &= \mathrm{M}(T(1{,}1{,}n-4), x) + x\,\mathrm{M}(T(1{,}1{,}n-6), x)\\ &< \mathrm{M}(T(1{,}2{,}n-5), x) + x\,\mathrm{M}(T(1{,}2{,}n-7), x) \text{ if } n \geq 7\\ &= \mathrm{M}(H(1{,}2{,}1{,}1{,}n), x),\end{aligned} \tag{4.5}$$

$$\begin{aligned}\mathrm{M}(H(1{,}2{,}1{,}1{,}n), x) &= \mathrm{M}(H(1{,}1{,}1{,}1{,}n-1), x) + x\,\mathrm{M}(T(1{,}1{,}n-5), x)\\ &< \mathrm{M}(H(1{,}1{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}2{,}n-6), x) \text{ if } n \geq 8\\ &= \mathrm{M}(H(1{,}2{,}1{,}2{,}n), x),\end{aligned} \tag{4.6}$$

$$\begin{aligned}\mathrm{M}(H(1{,}1{,}2{,}2{,}n), x) &= \mathrm{M}(H(1{,}1{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}1{,}n-5), x)\\ &< \mathrm{M}(H(1{,}1{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}2{,}n-6), x) \text{ if } n \geq 6\\ &= \mathrm{M}(H(1{,}2{,}1{,}2{,}n), x),\end{aligned} \tag{4.7}$$

$$\begin{aligned}\mathrm{M}(H(1{,}2{,}1{,}2{,}n), x) &= \mathrm{M}(H(1{,}1{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}2{,}n-6), x)\\ &< \mathrm{M}(H(1{,}2{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}2{,}n-6), x) \text{ if } n \geq 9\\ &= \mathrm{M}(H(1{,}2{,}2{,}2{,}n), x),\end{aligned} \tag{4.8}$$

$$\begin{aligned}\mathrm{M}(H(1{,}2{,}2{,}2{,}n), x) &= \mathrm{M}(H(1{,}2{,}1{,}2{,}n-1), x) + x\,\mathrm{M}(T(1{,}2{,}n-6), x)\\ &< \mathrm{M}(H(1{,}2{,}2{,}2{,}n-1), x) + x\,\mathrm{M}(T(2{,}2{,}n-7), x) \text{ if } n \geq 10\\ &= \mathrm{M}(H(2{,}2{,}2{,}2{,}n), x),\end{aligned} \tag{4.9}$$

$$\begin{aligned}
&\mathrm{M}(H(2,2,2,n-7,n),x)\\
&= \mathrm{M}(H(1,2,2,n-7,n-1),x) + x\,\mathrm{M}(T(2,2,n-7),x)\\
&= x\,\mathrm{M}(P_2,x)\,\mathrm{M}(P_2,x)\,\mathrm{M}(P_{n-7},x) + (1+x)\,\mathrm{M}(T(2,2,n-7),x)\\
&< x\,\mathrm{M}(P_2,x)\,\mathrm{M}(T(2,2,n-10),x) + (1+x)\,\mathrm{M}(T(2,2,n-7),x) \text{ for } n \geq 10\\
&= \mathrm{M}(H(2,2,2,2,n),x) \qquad (4.10)
\end{aligned}$$

and for $3 \leq d_{22} \leq n-9$ (and hence $n \geq 12$)

$$\begin{aligned}
&\mathrm{M}(H(2,2,2,d_{22},n),x)\\
&= (1+x)\,\mathrm{M}(T(2,d_{22},n-5-d_{22}),x)\\
&\qquad + x\,\mathrm{M}(P_2,x)\,\mathrm{M}(T(2,d_{22},n-8-d_{22}),x)\\
&< (1+x)\,\mathrm{M}(T(2,2,n-7),x) + x\,\mathrm{M}(P_2,x)\,\mathrm{M}(T(2,2,n-10,x)\\
&= \mathrm{M}(H(2,2,2,2,n),x). \qquad (4.11)
\end{aligned}$$

More inequalities are obtained by induction in the next two lemmas.

Lemma 4.5 *For all integers $n \geq 10$ and for all real numbers $x > 0$ we have*

$$\mathrm{M}(H(1,1,2,d_{22},n),x) < \mathrm{M}(H(2,2,2,2,n),x)$$

for any integer d_{22} such that $n-5 \geq d_{22} \geq 1$, and

$$\mathrm{M}(H(1,2,2,d_{22},n),x) < \mathrm{M}(H(2,2,2,2,n),x) \qquad (4.12)$$

for any integer $d_{22} \geq 1$ at most equal to $n-6$.

Proof. Induction with respect to d_{22}: The initial cases corresponding to $d_{22} \in \{1,2\}$ were already obtained in (4.6), (4.7), (4.8) and (4.9). Assume that the lemma is true whenever $1 \leq d_{22} \leq k$ for some integer $k \geq 2$. Now consider the case where $d_{22} = k+1$. Using the relations in (4.3) and (4.4) we obtain

$$\begin{aligned}
&\mathrm{M}(H(1,1,2,d_{22},n),x)\\
&\qquad = \mathrm{M}(H(1,1,2,k+1,n),x)\\
&\qquad = \mathrm{M}(H(1,1,2,k,n-1),x) + x\,\mathrm{M}(H(1,1,2,k-1,n-2),x)\\
&\qquad < \mathrm{M}(H(2,2,2,2,n-1),x) + x\,\mathrm{M}(H(2,2,2,2,n-2),x)\\
&\qquad = \mathrm{M}(H(2,2,2,2,n),x)
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{M}(H(1,2,2,d_{22},n),x)\\
&\qquad = \mathrm{M}(H(1,2,2,k+1,n),x)\\
&\qquad = \mathrm{M}(H(1,2,2,k,n-1),x) + x\,\mathrm{M}(H(1,2,2,k-1,n-2),x)\\
&\qquad < \mathrm{M}(H(2,2,2,2,n-1),x) + x\,\mathrm{M}(H(2,2,2,2,n-2),x)\\
&\qquad = \mathrm{M}(H(2,2,2,2,n),x).
\end{aligned}$$

□

Lemma 4.6 *Let* $n \geq 10$, $d_{22} \geq d_{12}$ *and* $n-5 \geq d_{12}+d_{22}$. *Then we have* $\mathrm{M}(H(2,d_{12},2,d_{22},n),x) < \mathrm{M}(H(2,2,2,2,n),x)$ *except if* (d_{12},d_{22}) *is equal to* $(2,2)$ *or* $(2,n-8)$.

Proof. If $d_{22} = n-8$, then we only have to check for $d_{12} = 1,3$: in $H(2,d_{12},2,d_{22},n)$ if $d_{12} \geq 4$, then $d_{22} < n-8$. The case of $d_{12} = 1$ follows from (4.12). In $H(2,3,2,n-8,n)$, the two vertices of degree 3 in Figure 4.4 coincide and become a single vertex of degree 4, hence using Lemma 4.1 and (4.10) we obtain

$$\mathrm{M}(H(2,3,2,n-8,n),x) < \mathrm{M}(H(2,2,2,n-7,n),x) < \mathrm{M}(H(2,2,2,2,n),x),$$

which covers the case of $d_{12} = 3$.

For any given value of $d_{22} \neq n-8$ we reason by induction with respect to d_{12}. The initial cases corresponding to $d_{12} = 1,2$ were already obtained in (4.10), (4.11) and (4.12). Note that $H(1,2,2,d_{22},n)$ and $H(2,1,2,d_{22},n),x)$ are isomorphic. The induction step follows from (4.3) and (4.4) in a similar way as in the proof of Lemma 4.5. $\square$

We are left to compare $\mathrm{En}(H(2,2,2,2,n))$ and $\mathrm{En}(H(2,2,2,n-8,n))$. As we will observe in the rest of this section the sign of

$$\mathrm{M}(H(2,2,2,2,n),x) - \mathrm{M}(H(2,2,2,n-8,n),x)$$

depends on x. Therefore, we have to estimate each of $\mathrm{En}(H(2,2,2,2,n))$ and $\mathrm{En}(H(2,2,2,n-8,n))$ in order to be able to compare them. For this we need explicit expressions for $\mathrm{M}(H(2,2,2,2,n),x^2)$ and $\mathrm{M}(H(2,2,2,n-8,n),x^2)$.

Using (4.3) we obtain the recurrence relation

$$\mathrm{M}(H(2,2,2,2,n+2),x^2) = \mathrm{M}(H(2,2,2,2,n+1),x^2) + x^2\,\mathrm{M}(H(2,2,2,2,n),x^2).$$

Its characteristic polynomial $P(t) = t^2 - t - x^2$ has two roots

$$t_1 = \frac{1+\sqrt{1+4x^2}}{2} = \frac{-1}{z^2-1} \text{ and } t_2 = \frac{1-\sqrt{1+4x^2}}{2} = \frac{z^2}{z^2-1},$$

where $x = z/(1-z^2)$. To have x ranging in $(0,+\infty)$ we take $0 < z < 1$. This implies that

$$\mathrm{M}(H(2,2,2,2,9+k),x^2) = A(z)\left(\frac{z^2}{z^2-1}\right)^k + B(z)\left(\frac{-1}{z^2-1}\right)^k \tag{4.13}$$

for some $A(z)$ and $B(z)$ which satisfy

$$\begin{cases} A(z)+B(z) = \mathrm{M}(H(2,2,2,2,9),x^2) \\ \qquad\qquad = \dfrac{(z^4-z^2+1)^3(z^4+3z^2+1)}{(z^2-1)^8} \\ A(z)\dfrac{z^2}{z^2-1}+B(z)\dfrac{-1}{z^2-1} = \mathrm{M}(H(2,2,2,2,10),x^2) \\ \qquad\qquad = \dfrac{(z^4-z^2+1)^2(z^{12}+z^{10}-2z^8+z^6-2z^4+z^2+1)}{(z^2-1)^{10}}. \end{cases} \tag{4.14}$$

Solving the system of equations we get

$$A(z) = \frac{z^4(z^4-z^2+1)^2(z^4+z^2-1)^2}{(z^2-1)^9(z^2+1)}$$

and

$$B(z) = -\frac{(z^4-z^2-1)^2(z^4-z^2+1)^2}{(z^2-1)^9(z^2+1)}.$$

Hence, (4.13) becomes

$$\begin{aligned} &\mathrm{M}(H(2,2,2,2,n),x^2) \\ &= \frac{(z^4-z^2+1)^2}{z^2+1}\left(z^{-14}(z^4+z^2-1)^2\left(\frac{z^2}{z^2-1}\right)^n+(z^4-z^2-1)^2\left(\frac{-1}{z^2-1}\right)^n\right) \\ &= \frac{(z^4-z^2+1)^2}{(z^2+1)(1-z^2)^n}\left(z^{-14}(z^4+z^2-1)^2(-1)^n z^{2n}+(z^4-z^2-1)^2\right). \end{aligned} \tag{4.15}$$

Solving the system of equations (4.14) where $\mathrm{M}(H(2,2,2,2,9),x^2)$ and $\mathrm{M}(H(2,2,2,2,10),x^2)$ are replaced by $\mathrm{M}(H(2,2,2,9-8,9),x^2)$ and $\mathrm{M}(H(2,2,2,10-8,10),x^2)$ respectively, we obtain

$$A(z) = \frac{z^6(z^4-z^2+1)(z^{10}+z^8-2z^6+2z^4-2z^2+1)}{(z^2-1)^9(z^2+1)}$$

and

$$B(z) = -\frac{(z^4-z^2+1)(z^{10}-2z^8+2z^6-2z^4+z^2+1)}{(z^2-1)^9(z^2+1)},$$

which lead to

$$\begin{aligned} &\mathrm{M}(H(2,2,2,n-8,n),x^2) \\ &= \frac{z^4-z^2+1}{(z^2+1)(1-z^2)^n}\big((-1)^n z^{2n-12}(z^{10}+z^8-2z^6+2z^4-2z^2+1) \\ &\qquad\qquad +z^{10}-2z^8+2z^6-2z^4+z^2+1\big). \end{aligned}$$

It is convenient to use the following abbreviations:

$$\begin{aligned}
Q_1(z) &= (z^4 - z^2 + 1)(z^4 + z^2 - 1)^2, \\
Q_2(z) &= z^{12} + z^{10} - 2z^8 + 2z^6 - 2z^4 + z^2 \\
&= z^{12} + (z^5 - z^3)^2 + (z^3 - z)^2, \\
R_1(z) &= (z^4 - z^2 + 1)(z^4 - z^2 - 1)^2, \\
R_2(z) &= z^{10} - 2z^8 + 2z^6 - 2z^4 + z^2 + 1.
\end{aligned}$$

Note that

$$\begin{aligned}
R_1(z) - R_2(z) &= -z^6(Q_1(z) - Q_2(z)) \\
&= (z^2 - 1)z^6(z^2 - z - 1)\left(z - \frac{\sqrt{5}-1}{2}\right)\left(z + \frac{\sqrt{5}+1}{2}\right). \qquad (4.16)
\end{aligned}$$

Equation (3.14) can be rewritten in terms of z as

$$\mathrm{En}(T) = \frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log \mathrm{M}(T, x^2)dz.$$

For even n we have

$$\begin{aligned}
&\frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)} \\
&\quad = \frac{z^{2n-14}Q_1(z) + R_1(z)}{z^{2n-14}Q_2(z) + R_2(z)} \\
&\quad = 1 + \frac{z^{2n-14}(Q_1(z) - Q_2(z)) + R_1(z) - R_2(z)}{z^{2n-14}Q_2(z) + R_2(z)} \\
&\quad = 1 + \frac{(R_1(z) - R_2(z))(1 - z^{2n-20})}{z^{2n-14}Q_2(z) + R_2(z)} \\
&\quad = 1 + \frac{(z^2-1)(z^2-z-1)\left(z - \frac{\sqrt{5}-1}{2}\right)\left(z + \frac{\sqrt{5}+1}{2}\right)(1 - z^{2n-20})z^6}{z^{2n-14}Q_2(z) + R_2(z)}.
\end{aligned}$$

Let

$$\begin{aligned}
&I_-(n) \\
&= \frac{2}{\pi}\int_0^{\frac{\sqrt{5}-1}{2}} \frac{1}{z^2}(1+z^2)\log\frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)}dz \\
&\geq \frac{2}{\pi}\int_0^{\frac{\sqrt{5}-1}{2}} \frac{1}{z^2}(1+z^2)\log\left(1 + \frac{(z^2-1)\left(z - \frac{\sqrt{5}-1}{2}\right)\left(z + \frac{\sqrt{5}+1}{2}\right)(z^2-z-1)z^6}{R_2(z)}\right)dz \\
&> -0.003
\end{aligned}$$

and for $n \geq 12$ let

$$
\begin{aligned}
I_+(n) &= \frac{2}{\pi} \int_{\frac{\sqrt{5}-1}{2}}^{1} \frac{1}{z^2}(1+z^2) \log \frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)} dz \\
&\geq \frac{2}{\pi} \int_{\frac{\sqrt{5}-1}{2}}^{1} \frac{1}{z^2}(1+z^2) \log \Bigg(1+ \\
&\qquad \frac{(z^2-1)\left(z-\frac{\sqrt{5}-1}{2}\right)\left(z+\frac{\sqrt{5}+1}{2}\right)(z^2-z-1)(1-z^{2\cdot 12-20})z^6}{z^{2\cdot 12-14}Q_2(z)+R_2(z)}\Bigg) dz \\
&> 0.009
\end{aligned}
$$

to end up with

$$
\mathrm{En}(H(2,2,2,2,n)) - \mathrm{En}(H(2,2,2,n-8,n)) = I_-(n) + I_+(n) > 0 \qquad (4.17)
$$

whenever n is even and at least 12.

For odd n we have

$$
\begin{aligned}
&\frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)} \\
&\quad = \frac{R_1(z) - z^{2n-14}Q_1(z)}{R_2(z) - z^{2n-14}Q_2(z)} \\
&\quad = 1 + \frac{(z^2-1)z^6(z^2-z-1)\left(z-\frac{\sqrt{5}-1}{2}\right)\left(z+\frac{\sqrt{5}+1}{2}\right)(1+z^{2n-20})}{R_2(z) - z^{2n-14}Q_2(z)}.
\end{aligned}
$$

Let

$$
\begin{aligned}
J_-(n) &= \frac{2}{\pi} \int_{0}^{\frac{\sqrt{5}-1}{2}} \frac{1}{z^2}(1+z^2) \frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)} dz \\
&\geq \frac{2}{\pi} \int_{0}^{\frac{\sqrt{5}-1}{2}} \frac{1}{z^2}(1+z^2) \log \Bigg(1+ \\
&\qquad \frac{(z^2-1)(z^2-z-1)\left(z-\frac{\sqrt{5}-1}{2}\right)\left(z+\frac{\sqrt{5}+1}{2}\right)(1+z^{2\cdot 11-20})z^6}{R_2(z) - z^{2\cdot 11-14}Q_2(z)}\Bigg) dz \\
&> -0.004
\end{aligned}
$$

and

$$
\begin{aligned}
&J_+(n)\\
&= \frac{2}{\pi} \int_{\frac{\sqrt{5}-1}{2}}^{1} \frac{1}{z^2}(1+z^2)\frac{\mathrm{M}(H(2,2,2,2,n),x^2)}{\mathrm{M}(H(2,2,2,n-8,n),x^2)}dz\\
&\geq \frac{2}{\pi} \int_{\frac{\sqrt{5}-1}{2}}^{1} \frac{1}{z^2}(1+z^2)\log\left(1+\frac{(z^2-1)(z^2-z-1)\left(z-\frac{\sqrt{5}-1}{2}\right)\left(z+\frac{\sqrt{5}-1}{2}\right)z^6}{R_2(z)}\right)dz\\
&> 0.021.
\end{aligned}
$$

Again this leads to

$$\mathrm{En}(H(2,2,2,2,n)) - \mathrm{En}(H(2,2,2,n-8,n)) = J_-(n) + J_+(n) > 0$$

for odd $n \geq 11$. The conclusion for this section is summarized in the following theorem (the cases of $n = 9, 10$ can be checked easily):

Theorem 4.7 *Among all trees with at least four leaves and order n at least 9, $H(2,2,2,2,n)$ is the unique tree with maximum energy.*

4.3 Comparison of $\mathrm{En}(H(2,2,2,2,n))$ with the energy of tripods

It will be convenient to use the following abbreviation:

$$g_{a,n,r}(i) := a^i + (-1)^r a^{n-i}. \tag{4.18}$$

For all non-negative integers n, r, $i < n/2$ and for any real number $a \in (0,1)$ we have

$$\frac{d}{di} g_{a,n,r}(i) = (a^i + (-1)^{r+1}a^{n-i})\log a \leq (a^i - a^{n-i})\log a < 0$$

and

$$g_{a,n,r}(i) \geq a^i - a^{n-i} > 0,$$

showing that $g_{a,n,r}$ is positive and decreasing under the above conditions.

For the tripod $T(i,j,k)$ of order n, we can assume $1 \leq i \leq j \leq k = n-i-j-1$ without loss of generality. Let v be the neighbor of the center of $T(i,j,k)$ in its branch of length k. Then we know that

$$\begin{aligned}
&\mathrm{M}(T(i,j,k),x^2)\\
&= \mathrm{M}(T(i,j,k)-v,x^2) + x^2 \sum_{w\in N(v)} \mathrm{M}(T(i,j,k)-\{v,w\},x^2)\\
&= \mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-1},x^2)\\
&+ x^2\,\mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-2},x^2) + x^2\,\mathrm{M}(P_i,x^2)\,\mathrm{M}(P_j,x^2)\,\mathrm{M}(P_{k-1},x^2). \qquad (4.19)
\end{aligned}$$

In a similar way as we got (4.15) we also obtain

$$\mathrm{M}(P_n,x^2) = \frac{z^2}{z^2+1}\left(\frac{z^2}{z^2-1}\right)^n + \frac{1}{z^2+1}\left(\frac{-1}{z^2-1}\right)^n. \qquad (4.20)$$

Using (4.20) and the relation $i+j+k=n-1$, after some straightforward calculations we obtain

$$\begin{aligned}
&x^2\,\mathrm{M}(P_i,x^2)\,\mathrm{M}(P_j,x^2)\,\mathrm{M}(P_{k-1},x^2)\\
&= \frac{1}{(z^2+1)^3(1-z^2)^n}\big((-1)^{n-2}z^{2(n+2)} + (-1)^{k-1}z^{2k+2} + (-1)^{j+k-1}z^{2(j+k+2)}\\
&+(-1)^{i+k-1}z^{2(i+k+2)} + (-1)^{i+j}z^{2(i+j+3)} + z^2 + (-1)^j z^{2(j+2)} + (-1)^i z^{2(i+2)}\big),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-1},x^2)\\
&= \frac{1}{(z^2+1)^2(1-z^2)^{n-1}}\big((-1)^{n-1}z^{2(n+1)} + 1 - (-1)^{i+j}z^{2(n-k+1)} + (-1)^{k+1}z^{2k}\big),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-2},x^2)\\
&= \frac{1}{(z^2+1)^2(1-z^2)^{n-2}}\big((-1)^n z^{2n} + 1 - (-1)^{i+j}z^{2(n-k+1)} + (-1)^k z^{2(k-1)}\big)
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-1},x^2) + x^2\,\mathrm{M}(P_{i+j+1},x^2)\,\mathrm{M}(P_{k-2},x^2)\\
&= \frac{1}{(z^2+1)^2(1-z^2)^n}\big((-1)^n z^{2(n+2)} + 1 - (-1)^{i+j}z^{2(n-k+1)} + (-1)^k z^{2(k+1)}\big).
\end{aligned}$$

Consequently, (4.19) becomes

$$\begin{aligned}
&\mathrm{M}(T(i,j,n-i-j-1),x^2)\\
&= \frac{1}{(z^2+1)^3(1-z^2)^n}\big((-1)^n z^{2(n+2)} + 1 - (-1)^{i+j} z^{2(n-k+1)} + (-1)^k z^{2(k+1)}\\
&+ (-1)^n z^{2(n+3)} + z^2 - (-1)^{i+j} z^{2(n-k+2)} + (-1)^k z^{2(k+2)} + (-1)^{n-2} z^{2(n+2)}\\
&+ (-1)^{k-1} z^{2k+2} + (-1)^{j+k-1} z^{2(j+k+2)} + (-1)^{i+k-1} z^{2(i+k+2)} + (-1)^{i+j} z^{2(i+j+3)}\\
&+z^2 + (-1)^j z^{2(j+2)} + (-1)^i z^{2(i+2)}\big)\\
&= \frac{1}{(z^2+1)^3(1-z^2)^n}\big((-1)^n z^{2(n+2)}(2+z^2) + 1 + 2z^2 - (-1)^{i+j} z^{2(i+j+2)}\\
&- (-1)^{n-(i+j)} z^{2(n+3-(i+j+2))} + (-1)^{n-i} z^{2(n+3-i)} + (-1)^{n-j} z^{2(n+3-j)}\\
&+(-1)^j z^{2(j+2)} + (-1)^i z^{2(i+2)}\big)\\
&= \frac{1}{(z^2+1)^3(1-z^2)^n}\big((-1)^n z^{2(n+2)}(2+z^2) + 1 + 2z^2 + (-1)^i g_{z^2,n+3,n}(i+2)\\
&\qquad +(-1)^j g_{z^2,n+3,n}(j+2) - (-1)^{i+j} g_{z^2,n+3,n}(i+j+2)\big)\,. \qquad (4.21)
\end{aligned}$$

Using the expressions in (4.15) and (4.21) we get

$$\begin{aligned}
&D(i,j,n,z)\\
&:= \frac{\mathrm{M}(T(i,j,n-1-i-j),x^2)}{\mathrm{M}(H(2,2,2,2,n),x^2)}\\
&= \frac{(-1)^n z^{2(n+2)}(2+z^2) + 1 + 2z^2 + (-1)^i g_{z^2,n+3,n}(i+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2(-1)^n z^{2n} + (z^4-z^2-1)^2\right)}\\
&+ \frac{(-1)^j g_{z^2,n+3,n}(j+2) - (-1)^{i+j} g_{z^2,n+3,n}(i+j+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2(-1)^n z^{2n} + (z^4-z^2-1)^2\right)}.
\end{aligned} \qquad (4.22)$$

Let us show that each of the four sequences $(D(i,j,n=2m,z))_{5\le m\in\mathbb{N}}$, $(D(i,j,2n,z))_{i\in\mathbb{N}}$, $(D(i,j,n=2m+1,z))_{4\le m\in\mathbb{N}}$ and $(D(i,j,2n+1,z))_{i\in\mathbb{N}}$ is dominated by an integrable function, which means that we will be able to interchange limits and integrals. The inequality $n-1-i-j\ge j\ge i\ge 1$ will be used. Note that for a function $f(n) = az^n/(bz^n+c)$ which has no pole in $(0,\infty)$, the derivative in $(0,\infty)$ is $f'(n) = (acz^n\log z)/(bz^n+c)^2$. In the case where $z\in(0,1)$, the function is increasing if $ac<0$ and decreasing if $ac>0$.

For all $n\ge 9$ we deduce from (4.22) that

$$\begin{aligned}
&D(i,j,n,z)\\
&\le \frac{z^{2(n+2)}(2+z^2) + 1 + 2z^2 + g_{z^2,n+3,2}(1+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2 z^{2n} + (z^4-z^2-1)^2\right)}\\
&+ \frac{g_{z^2,n+3,2}(1+2) + g_{z^2,n+3,2}(1+1+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2 z^{2n} + (z^4-z^2-1)^2\right)}
\end{aligned}$$

$$= \frac{1+2z^2+2z^6+z^8+z^{2n}(2z^4+z^6+2+z^{-2})}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$\leq e_1(z) := \frac{1+2z^2+2z^6+z^8+z^{18}(2z^4+z^6+2+z^{-2})}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2z^{18}+(z^4-z^2-1)^2\right)},$$

for even $n \geq 10$ we have

$$D(i,j,n,z)$$
$$\geq \frac{z^{2(n+2)}(2+z^2)+1+2z^2-g_{z^2,n+3,2}(1+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$+\frac{-g_{z^2,n+3,2}(1+2)-g_{z^2,n+3,2}(2+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$= \frac{1+2z^2-2z^6-z^8+z^{2n}(2z^4+z^6-2-z^{-2})}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$\geq e_2(z) := \frac{1+2z^2-2z^6-z^8+z^{20}(2z^4+z^6-2-z^{-2})}{(1+z^2)^2(z^4-z^2+1)^2\left(z^{-14}(z^4+z^2-1)^2z^{20}+(z^4-z^2-1)^2\right)},$$

for odd $n \geq 9$ we get

$$D(i,j,n,z)$$
$$\geq \frac{-z^{2(n+2)}(2+z^2)+1+2z^2-g_{z^2,n+3,1}(1+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$+\frac{-g_{z^2,n+3,1}(1+2)-g_{z^2,n+3,1}(2+2)}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$= \frac{1+2z^2-2z^6-z^8+z^{2n}(-2z^4-z^6+2+z^{-2})}{(1+z^2)^2(z^4-z^2+1)^2\left(-z^{-14}(z^4+z^2-1)^2z^{2n}+(z^4-z^2-1)^2\right)}$$
$$\geq e_3(z) := \frac{1+2z^2-2z^6-z^8}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}$$

and consequently

$$\left|\frac{1+z^2}{z^2}\log D(i,j,n,z)\right| \leq e_4(z)$$
$$:= \max\left\{0, \frac{1+z^2}{z^2}\log e_1(z)\right\} - \min\left\{0, \frac{1+z^2}{z^2}\log e_2(z), \frac{1+z^2}{z^2}\log e_3(z)\right\}.$$

In view of

$$-0.92 \leq \int_0^1 \frac{1}{z^2}(1+z^2)\log e_2(z)dx \leq \int_0^1 \frac{1}{z^2}(1+z^2)\log e_3(z)dx$$
$$\leq \int_0^1 \frac{1}{z^2}(1+z^2)\log e_1(z)dx \leq 0.61$$

we can deduce that the function e_4 is integrable in $(0,1)$. These observations will allow us to interchange limits and integrals using Lebesgue's dominated convergence theorem.

Note that under the assumption that $i \le j \le k$ and $i+j+k = n-1$, if i tends to infinity, then necessarily $j, k, n-i, n-j, n-k$ also tend to infinity, hence we have

$$\begin{aligned}\lim_{i\to\infty} D(i,j,n,z) &= \frac{1+2z^2}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}\\ &= \frac{1+2z^2}{(1+z^2)^2((z^4-z^2)^2-1)^2}\end{aligned}$$

and

$$\begin{aligned}D(i,j,\infty,z) &:= \lim_{n\to\infty} D(i,j,n,z)\\ &= \frac{1+2z^2+(-1)^i z^{2i+4}+(-1)^j z^{2j+4}-(-1)^{i+j}z^{2i+2j+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}. \qquad (4.23)\end{aligned}$$

Therefore we get (remember the relation $x = z/(1-z^2)$)

$$\begin{aligned}&\lim_{i\to\infty} \mathrm{En}(T(i,j,n-i-j-1)) - \mathrm{En}(H(2,2,2,2,n))\\ &= \lim_{i\to\infty} \frac{2}{\pi}\int_0^\infty \frac{1}{x^2}\log\frac{\mathrm{M}(T(i,j,n-1-i-j),x^2)}{\mathrm{M}(H(2,2,2,2,n),x^2)}dx\\ &= \frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log\frac{1+2z^2}{(1+z^2)^2((z^4-z^2)^2-1)^2}dz < -0.014. \qquad (4.24)\end{aligned}$$

This shows that there are only finitely many values of i for which the energy of $T(i,j,n-i-j-1)$ can be greater than that of $H(2,2,2,2,n)$. Next we determine all such values of i.

Lemma 4.8 *For n large enough, if*

$$\mathrm{En}(T(i,j,n-i-j-1)) > \mathrm{En}(H(2,2,2,2,n)),$$

then $i \in I = \{1,2,3,4,5,6,8,10,12,14,16,18\}$.

Proof. We use the notation in (4.23).

$a)$ For even $i = 2k$ and even $j = 2(k+l)$ we obtain:

$$\begin{aligned}D(i,j,\infty,z) = ee(i,j,z) &:= \frac{1+2z^2+z^{2i+4}+z^{2j+4}-z^{2i+2j+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}\\ &\le ee(i,i,z) = \frac{1+2z^2+2z^{2i+4}-z^{4i+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}\\ &\le ee(20,20,z) \text{ for all } i \ge 20,\end{aligned}$$

and

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log ee(20,20,z)dz < -0.001.$$

This shows that for n large enough, $k \geq 10$ and $l \geq 0$ we have $\mathrm{En}(T(2k, 2(k+l), n-4k-2l-1)) < \mathrm{En}(H(2,2,2,2,n))$.

$b)$ For even $i = 2k$ and odd $j = 2(k+l)+1$ we obtain:

$$\begin{aligned} D(i,j,\infty,z) &= eo(i,j,z) \\ &:= \frac{1+2z^2+z^{2i+4}-z^{2j+4}+z^{2i+2j+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\ &\leq eo(i,\infty,z) = \frac{1+2z^2+z^{2i+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\ &\leq eo(14,\infty,z) \text{ for all } i \geq 14, \end{aligned}$$

and

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log eo(14,\infty,z)dz < -0.001.$$

This means that (for n large enough and $l \geq 0$) the energy of a tripod $T(2k, 2(k+l)+1, n-4k-2l-2)$ can only be greater than that of $H(2,2,2,2,n)$ if $k \leq 6$.

$c)$ For odd $i = 2k+1$ and even $j = 2(k+l+1)$ we obtain:

$$\begin{aligned} D(i,j,\infty,z) &= \frac{1+2z^2-z^{2i+4}+z^{2j+4}+z^{2i+2j+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\ &\leq \frac{1+2z^2-z^{2i+4}+z^{2i+6}+z^{4i+6}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\ &\leq oe(i,z) := \frac{1+2z^2+z^{4i+6}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\ &\leq oe(7,z) \text{ for all } i \geq 7 \end{aligned}$$

and

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log oe(7,z)dz < -0.002.$$

Hence (for n large enough) for all integers $l \geq 0$ a tripod $T(2k+1, 2(k+l+1), n-4k-2l-4)$ that can possibly have greater energy than that of $H(2,2,2,2,n)$ must satisfy $k \in \{0,1,2\}$.

$d)$ For odd i and odd j we obtain:

$$
\begin{aligned}
D(i,j,\infty,z) = oo(i,j,z) &:= \frac{1+2z^2-z^{2i+4}-z^{2j+4}-z^{2i+2j+4}}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2} \\
&\leq oo(\infty,\infty,z) = \frac{1+2z^2}{(1+z^2)^2(z^4-z^2+1)^2(z^4-z^2-1)^2}
\end{aligned}
$$

where as we have seen in (4.24)

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log oo(\infty,\infty,z)dz < -0.014.$$

□

For any given value of i, Theorem 4.3 allows us to obtain the complete list of all tripods of order n and with shortest branch of length i, ordered by their energies. In the following we determine the place of $H(2,2,2,2,n)$ in each list corresponding to a value in I. For $i = 1$ we have

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log D(1,2,\infty,z)dz > 0.004$$

and

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log D(1,4,\infty,z)dz < -0.034,$$

thus

$$
\begin{aligned}
\mathrm{En}(T(1,2,n-4)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(1,4,n-6)) \\
> \cdots > \mathrm{En}(T(1,1,n-3)),
\end{aligned}
$$

if n is large enough. Since

$$\frac{2}{\pi}\int_0^1 \frac{1}{z^2}(1+z^2)\log D(2,3,\infty,z)dz > 0.030$$

we deduce that for $i = 2$ and n large enough we have

$$\mathrm{En}(T(2,2,n-5)) > \cdots > \mathrm{En}(T(2,3,n-6)) > \mathrm{En}(H(2,2,2,2,n)).$$

By similar arguments, for large enough n we also have:

- $\mathrm{En}(T(3,4,n-8)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(3,6,n-10)) > \cdots > \mathrm{En}(T(3,3,n-7))$,

- $\mathrm{En}(T(4,4,n-9)) > \cdots > \mathrm{En}(T(4,5,n-10)) > \mathrm{En}(H(2,2,2,2,n))$,

- $\mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(5,6,n-12)) > \cdots > \mathrm{En}(T(5,5,n-11))$,

- $\mathrm{En}(T(6,6,n-13)) > \cdots > \mathrm{En}(T(6,7,n-14)) > \mathrm{En}(H(2,2,2,2,n))$,

- $\mathrm{En}(T(8,8,n-17)) > \cdots > \mathrm{En}(T(8,11,n-20)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(8,9,n-18))$,

- $\mathrm{En}(T(10,10,n-21)) > \cdots > \mathrm{En}(T(10,21,n-32)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(10,19,n-30)) > \cdots > \mathrm{En}(T(10,11,n-22))$,

- $\mathrm{En}(T(12,12,n-25)) > \cdots > \mathrm{En}(T(12,85,n-98)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(12,83,n-96)) > \cdots > \mathrm{En}(T(12,13,n-26))$,

- $\mathrm{En}(T(14,14,n-29)) > \cdots > \mathrm{En}(T(14,30,n-45)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(14,32,n-47)) > \cdots > \mathrm{En}(T(14,15,n-30))$,

- $\mathrm{En}(T(16,16,n-33)) > \cdots > \mathrm{En}(T(16,22,n-49)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(16,24,n-41)) > \cdots > \mathrm{En}(T(16,17,n-34))$,

- $\mathrm{En}(T(18,18,n-37)) > \mathrm{En}(H(2,2,2,2,n)) > \mathrm{En}(T(18,20,n-39)) > \cdots > \mathrm{En}(T(18,19,n-38))$.

Knowing these, we can now count the tripods whose energies are greater than $\mathrm{En}(H(2,2,2,2,n))$ and obtain the following theorem:

Theorem 4.9 *For large enough n the quadripod $H(2,2,2,2,n)$ is the tree with $(3n-84)^{th}$ (resp. $(3n-87)^{th}$) largest energy for odd n (resp. for even n).*

Proof. To form an n-vertex tripod whose shortest branch has length i, we can merge the end of P_{i+1} with the j^{th} vertex in P_{n-i}, where $i+1 \leq j \leq \lfloor (n-i+1)/2 \rfloor$. This gives $\lfloor (n-i+1)/2 \rfloor - i$ possible tripods. For n large enough, if $\eta_{i,n}$ is the number of n-vertex tripods which have shortest branch of length i and larger energy than $H(2,2,2,2,n)$, then by Lemma 4.8 and the above inequalities we have

$$\begin{aligned}\sum_{i\in\mathbb{N}} \eta_{i,n} &= \eta_{1,n} + \eta_{3,n} + \eta_{5,n} + \eta_{14,n} + \eta_{16,n} + \eta_{18,n} + \sum_{l=1}^{6} \eta_{2l,n} \\ &= 1+1+0+9+4+1-42+\sum_{l=1}^{6} \left\lfloor \frac{n-2l+1}{2} \right\rfloor - 2l \\ &= -26 + 6\left\lfloor \frac{n+1}{2} \right\rfloor - 3\sum_{l=1}^{6} l \\ &= \begin{cases} 3n-86 \text{ if } n \text{ is odd,} \\ 3n-89 \text{ if } n \text{ is even.} \end{cases}\end{aligned}$$

Including the path, we obtain the theorem. □

At least for sufficiently large number of vertices, if a class of fixed order graphs contains a tripod with short enough even length branch, then it is very likely that this tripod has the largest energy in the class. The following theorem is an example of such a situation.

Theorem 4.10 *For all i in $J = \{1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 18\}$ and n large enough, the n-vertex tree with diameter $n-i-1$ and maximum energy is $T(i, 2\lceil i/2\rceil, n-2\lceil i/2\rceil - i - 1)$.*

Proof. Let i be an element of J and T_i be a tree of diameter $n-i-1$ which is maximal with respect to the energy. We know that $\operatorname{diam}(P_n) = n-1 > n-i-1$, hence $T_i \neq P_n$. As we have seen above for large enough n (in particular we assume $n \geq 3i+1$), there exists a tripod $T(i, j_0, n-j_0-i-1)$ which has diameter $n-i-1$ such that $\operatorname{En}(T(i, j_0, n-j_0-i-1)) > \operatorname{En}(H(2,2,2,2,n))$. Using Theorem 4.7 this implies that T_i is a tripod. More precisely $T_i = T(i, j, n-j-i-1)$ for some $j \geq i$, in order to satisfy $\operatorname{diam}(T_i) = n-i-1$. From Theorem 4.3, we get that if $j \neq 2\lceil i/2\rceil$, we have

$$\operatorname{En}(T(i, j, n-j-i-1)) < \operatorname{En}(T(i, 2\lceil i/2\rceil, n-2\lceil i/2\rceil - i - 1)).$$

Therefore, we conclude that $T_i = T(i, 2\lceil i/2\rceil, n-2\lceil i/2\rceil - i - 1)$. □

4.4 List of large energy trees

For simplicity we write $G > G'$ instead of $\operatorname{En}(G) > \operatorname{En}(G')$. By ordering all the tripods with larger energy than that of $H(2,2,2,2,n)$ we obtain the head of the list of trees ordered by decreasing energy, until the first appearance of a non-tripod. Each "..." in the list refers to the chain obtained for a fixed shortest branch by using Theorem 4.3. For any inequality that cannot be obtained from Theorem 4.3, see the values in the Appendix A. For the sake of simplicity, we shorten the notation for tripods, instead of $T(i, j, n-i-j-1)$ we simply write $T_n(i, j)$ for the n-vertex tripod with shortest branches of length i and j.

Theorem 4.11 *The head of the list of all trees ordered by decreasing energy is given as follows for large enough n:*

$$
\begin{array}{llllllllll}
P_n & > & T_n(2,2) & > & \cdots & > & T_n(2,7) & > & T_n(4,4) & > \\
T_n(2,5) & > & T_n(4,6) & > & T_n(2,3) & > & T_n(4,8) & > & \cdots & > \\
T_n(4,18) & > & T_n(6,6) & > & T_n(4,20) & > & \cdots & > & T_n(4,15) & > \\
T_n(6,8) & > & T_n(4,13) & > & T_n(4,11) & > & T_n(6,10) & > & T_n(4,9) & > \\
T_n(6,12) & > & T_n(8,8) & > & T_n(6,14) & > & T_n(4,7) & > & T_n(6,16) & > \\
T_n(6,18) & > & \cdots & > & T_n(6,26) & > & T_n(8,10) & > & T_n(6,28) & > \\
\cdots & > & T_n(6,39) & > & T_n(8,12) & > & T_n(6,37) & > & \cdots & > \\
T_n(6,23) & > & T_n(8,14) & > & T_n(10,10) & > & T_n(6,21) & > & T_n(4,5) & > \\
T_n(6,19) & > & T_n(8,16) & > & T_n(6,17) & > & T_n(6,15) & > & T_n(8,18) & > \\
T_n(8,20) & > & T_n(10,12) & > & T_n(8,22) & > & T_n(6,13) & > & T_n(8,24) & > \\
\cdots & > & T_n(8,30) & > & T_n(10,14) & > & T_n(8,32) & > & T_n(8,34) & > \\
T_n(8,36) & > & T_n(6,11) & > & T_n(8,38) & > & \cdots & > & T_n(8,56) & > \\
T_n(12,12) & > & T_n(8,58) & > & \cdots & > & T_n(8,86) & > & T_n(10,16) & > \\
T_n(8,88) & > & \cdots & > & T_n(8,49) & > & T_n(10,18) & > & T_n(8,47) & > \\
\cdots & > & T_n(8,33) & > & T_n(12,14) & > & T_n(10,20) & > & T_n(6,9) & > \\
T_n(8,31) & > & T_n(8,29) & > & T_n(8,27) & > & T_n(10,22) & > & T_n(8,25) & > \\
T_n(10,24) & > & T_n(8,23) & > & T_n(12,16) & > & T_n(10,26) & > & T_n(1,2) & > \\
T_n(8,21) & > & T_n(10,28) & > & T_n(10,30) & > & T_n(14,14) & > & T_n(10,32) & > \\
T_n(8,19) & > & T_n(10,34) & > & T_n(12,18) & > & T_n(10,36) & > & \cdots & > \\
T_n(10,44) & > & T_n(8,17) & > & T_n(10,46) & > & \cdots & > & T_n(10,52) & > \\
T_n(12,20) & > & T_n(10,54) & > & \cdots & > & T_n(10,70) & > & T_n(14,16) & > \\
T_n(10,72) & > & \cdots & > & T_n(10,182) & > & T_n(12,22) & > & T_n(10,184) & > \\
\cdots & > & T_n(10,175) & > & T_n(8,15) & > & T_n(10,173) & > & \cdots & > \\
T_n(10,69) & > & T_n(6,7) & > & T_n(12,24) & > & T_n(10,67) & > & \cdots & > \\
T_n(10,53) & > & T_n(14,18) & > & T_n(10,51) & > & T_n(10,49) & > & T_n(12,26) & > \\
T_n(10,47) & > & \cdots & > & T_n(10,41) & > & T_n(16,16) & > & T_n(12,28) & > \\
T_n(10,39) & > & T_n(10,37) & > & T_n(8,13) & > & T_n(12,30) & > & T_n(10,35) & > \\
T_n(14,20) & > & T_n(10,33) & > & T_n(12,32) & > & T_n(10,31) & > & T_n(12,34) & > \\
T_n(12,36) & > & T_n(10,29) & > & T_n(12,38) & > & T_n(14,22) & > & T_n(16,18) & > \\
T_n(12,40) & > & T_n(10,27) & > & T_n(12,42) & > & T_n(12,44) & > & T_n(12,46) & > \\
T_n(10,25) & > & T_n(12,48) & > & T_n(14,24) & > & T_n(12,50) & > & \cdots & > \\
T_n(12,64) & > & T_n(10,23) & > & T_n(12,66) & > & \cdots & > & T_n(12,70) & > \\
T_n(14,26) & > & T_n(16,20) & > & T_n(12,72) & > & \cdots & > & T_n(12,92) & > \\
T_n(8,11) & > & T_n(12,94) & > & \cdots & > & T_n(12,130) & > & T_n(18,18) & > \\
T_n(12,132) & > & \cdots & > & T_n(12,162) & > & T_n(14,28) & > & T_n(12,164) & >
\end{array}
$$

$$\begin{array}{llllllllll}
\cdots & > & T_n(12,224) > & T_n(10,21) & > & T_n(12,226) > & \cdots & > \\
T_n(12,219) > & T_n(3,4) & > & T_n(12,217) > & \cdots & > & T_n(12,111) > \\
T_n(14,30) & > & T_n(12,109) > & \cdots & > & T_n(12,99) & > & T_n(16,22) & > \\
T_n(12,97) & > & \cdots & > & T_n(12,85) & > & H(2,2,2,2,n). &
\end{array}$$

Seeing the published article corresponding to the above theorem, Shan, Shao, Zhang and He investigated how large n should be for the theorem to hold. They proved [76] that Theorem 4.11 is true for $n \geq 7526$.

Seeing that Theorem 4.11 only holds for very large n while many important hydrocarbon molecules have fewer than 100 carbon atoms, mathematical chemist Gutman suggested to study the case $n \leq 100$. It is well-known that the first four entries in the list Theorem 4.11 are the four largest energy trees for $n \geq 14$. Together with Gutman and his close collaborators Furtula and Cvetić, we checked how far the list is still valid for $n \in [10, 100]$, see [33]. Just as in [33], let us refer to the list of trees in Theorem 4.11 as the $\mathcal{A}$-ordering. Remember that the notation $T_n(i,j)$ assumes that $i \leq j \leq n-1-i-j$. In the $\mathcal{A}$-ordering corresponding to small n, if $n-1-i-j < j$ then the entry $T_n(i,j)$ is just ignored. For a given n, we proceed as follows:

i) First, we determine the elements of the set $\mathbb{T}_n$ of tripods with larger energy than $H(2,2,2,2,n)$. This is done by numerical integration using formula (4.22). We only consider n at least 9 so that $H(2,2,2,2,n)$ exists and we can use Theorem 4.7.

ii) Then, we only need to sort the elements of $\mathbb{T}_n$, by decreasing energy to obtain the list of n-vertex trees with largest energy until the first non-tripod tree. Adding P_n at the top and $H(2,2,2,2,n)$ at the bottom of the list, at this stage we get the list of the $\Lambda(n)$ largest energy trees, for some integer $\Lambda(n)$ such that the number of trees with larger energy than $H(2,2,2,2,n)$ is $\Lambda(n)-1$.

iii) The last step is to find out the number $\Omega(n)$ of the entries in the part of the list which agree with the list in Theorem 4.11. We remark that the way tripods are counted here is slightly different from our paper [33], so the values are slightly different.

Figure 4.5 summarizes the behavior of $\Lambda(n)$ and $\Omega(n)$ for n ranging from 9 to 250. For precise values of Ω and Λ see Appendix B. Both Ω and Λ grow almost linearly with n, with some jumps at specific values of n.

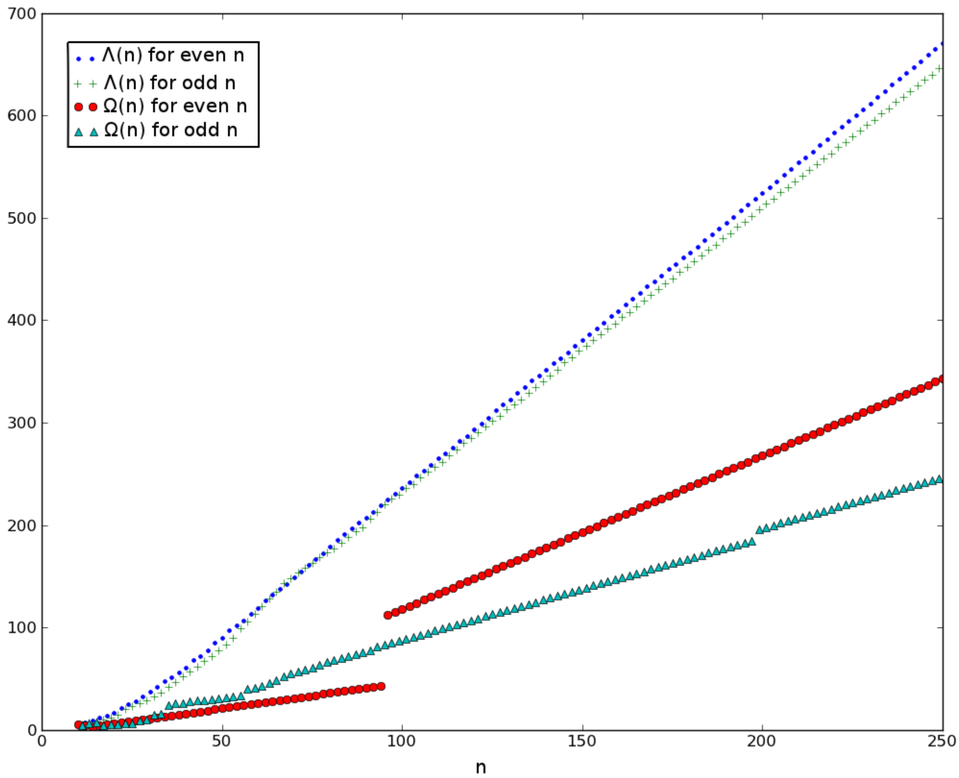


Figure 4.5: Dependence of the parameters Λ and Ω on the number of vertices n

Chapter 5

Energy, Hosoya index and Merrifield-Simmons index of trees with prescribed degree sequence

In this chapter, we present results which are a natural continuation of the works in [5, 42, 43]. In [42], the class $\mathbb{T}_{1,d}$ of trees whose degrees are either 1 or $d \geq 2$ was studied; the trees in $\mathbb{T}_{1,d}$ with largest and second-largest Hosoya index and energy and smallest/second-smallest Merrifield-Simmons index are characterized for any possible number of vertices. The two papers [5, 43] provide a characterization of the n-vertex tree whose maximum degree is $d+1$ which has maximum Merrifield-Simmons index and minimum energy and Hosoya index, for all positive integers n and d.

Definition 5.1 For a tree T of order n whose vertex degrees are $d_1 \geq \cdots \geq d_n$, the n-tuple $(d_1, \ldots, d_n)$ is called *degree sequence* of T.

We consider the class $\mathbb{T}_D$ of all trees which have a given degree sequence D. We find that for any degree sequence D, there exists a tree $\mathcal{M}(D) \in \mathbb{T}_D$ such that whenever $T \in \mathbb{T}_D$ either T and $\mathcal{M}(D)$ are isomorphic or the three inequalities $\mathrm{En}(\mathcal{M}(D)) < \mathrm{En}(T)$, $\mathrm{Z}(\mathcal{M}(D)) < \mathrm{Z}(T)$ and $\sigma(\mathcal{M}(D)) > \sigma(T)$ hold. Sections 5.1, 5.2, 5.3 are devoted to proving this observation and to describing the construction of $\mathcal{M}(D)$. As an additional result we show in Section 5.4 that if $B = (b_1, \ldots, b_n)$ and $D = (d_1, \ldots, d_n)$ are two different degree sequences such that

$$\sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} d_i$$

for all $1 \leq k \leq n$, then we have $\sigma(\mathcal{M}(B)) < \sigma(\mathcal{M}(D))$, $\mathrm{Z}(\mathcal{M}(B)) > \mathrm{Z}(\mathcal{M}(D))$ and $\mathrm{En}(\mathcal{M}(B)) > \mathrm{En}(\mathcal{M}(D))$. Several older results, which follow directly from this, will be revisited as applications of our main theorem.

Trees with prescribed degree sequence have also been studied in the context of other graph invariants, such as the spectral radius [10, 91] and Wiener-type graph invariants [75, 84, 85].

5.1 Preliminaries

In addition to definitions and a review of some lemmas from the literature, we already provide key lemmas towards the end of this introductory section.

Definition 5.2 If $(d_1, \ldots, d_n, 1, \ldots, 1)$ is the degree sequence of a tree T, where $d_n \geq 2$, then we call the n-tuple $(d_1, \ldots, d_n)$ *reduced degree sequence* of T.

If two trees T and T' have reduced degree sequence $(d_1, \ldots, d_n)$, and k and k' are respectively their numbers of leaves, then by the Handshake lemma we have

$$k + \sum_{j=1}^{n} d_j = 2(n + k - 1) \qquad \text{and} \qquad k' + \sum_{j=1}^{n} d_j = 2(n + k' - 1)$$

and hence

$$k = 2(1 - n) + \sum_{j=1}^{n} d_j = k'.$$

This shows that two trees with the same reduced degree sequence have the same degree sequence.

Definition 5.3 We call a subtree B of a tree T a *complete branch* of T if and only if $T - V(B)$ is connected. This means that T can be decomposed as in Figure 5.1 (a), where B and $T - B$ are non-empty.

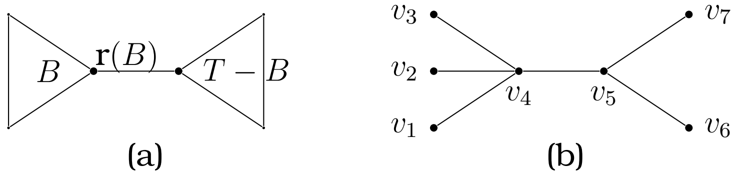


Figure 5.1: Example of complete branches

For example, if T is the tree in Figure 5.1 (b), then the subgraph spanned by $\{v_1, v_2, v_3, v_4\}$ is a complete branch, but the subgraph spanned by $\{v_1, v_2, v_4\}$ is not. For any complete branch B of T, we define its root to be the unique vertex which has a neighbor in $T - B$. We will denote by $\mathrm{rn}(B)$ the neighbor of $\mathrm{r}(B)$ in $T - B$, and $\mathrm{rd}(B)$ the degree of $\mathrm{r}(B)$ as a vertex of B. If $B_1, \dots, B_{\mathrm{rd}(B)}$ are the connected components of $B - \mathrm{r}(B)$, then we write $B = [B_1, \dots, B_{\mathrm{rd}(B)}]$.

For any two rooted trees R and R', we write $R \approx_r R'$ if and only if there exists an isomorphism $R \to R'$ which preserves the roots, otherwise we write $R \not\approx_r R'$.

Definition 5.4 We call a vertex in a tree a *pseudo-leaf* if and only if it is not a leaf and it has at most one neighbor with degree greater than 1. We call a complete branch whose root is a pseudo-leaf a *pseudo-leaf branch.*

We denote by $[d]$ a pseudo-leaf branch with d vertices. For any forest F, the set of the pseudo-leaves in F is denoted by $\mathcal{P}(F)$.

We adopt similar notations as in [43]. For every complete branch B of a tree, we define $\mathrm{m}_0(B, k)$ to be the number of matchings of cardinality k in B not covering $\mathrm{r}(B)$, $\mathrm{M}_0(B, x) = \sum_{k \geq 0} \mathrm{m}_0(B, k) x^k$ and

$$\tau(B, x) = \frac{\mathrm{M}_0(B, x)}{\mathrm{M}(B, x)}. \tag{5.1}$$

Furthermore, the next three lemmas from [43] will play important roles:

Lemma 5.5 ([43]) *Let $B = [B_1, \dots, B_{\mathrm{rd}(B)}]$ be a complete branch of a tree. Then, for all positive x we have*

$$\tau(B, x) = \frac{1}{1 + x \sum_{i=1}^{\mathrm{rd}(B)} \tau(B_i, x)}. \tag{5.2}$$

It is convenient to set $\tau(\varnothing, x) = 0$ for all $x > 0$, so that recurrence (5.2) *still holds if some of the B_i's are empty.*

Lemma 5.6 ([43]) *Let B be a complete branch of a tree and $x > 0$. Then*

$$\frac{1}{x\,\mathrm{rd}(B) + 1} \leq \tau(B, x) \leq 1.$$

Remark 5.7 Note that the upper bound 1 is reached only if B is a leaf. It follows from this that the lower bound is also obtained only if B is a pseudo-leaf branch.

We say that T is an x-*minimal* tree if $\mathrm{M}(T,x) \leq \mathrm{M}(T',x)$ for all trees T' with the same degree sequence as T. The exchange lemma used by Heuberger and Wagner in [43] is also going to play a key role in this chapter, it reads as follows:

Lemma 5.8 ([43]) *Let $x > 0$ and let T be a minimal tree with respect to $\mathrm{M}(.,x)$ among all trees with a given order and maximum degree d. If there are (possibly empty) rooted trees $L_1,\ldots,L_d,R_1,\ldots,R_d$ and a tree T_0 such that T can be decomposed as in Figure 5.2 and such*

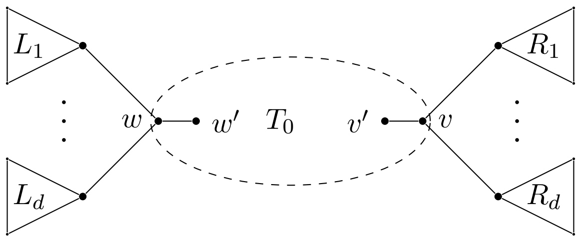


Figure 5.2: Decomposition of the tree T in Lemma 5.8

that $\tau(L_1,x) < \tau(R_1,x)$ (after appropriate reordering of the L_i's and the R_i's), then

$$\max\{\tau(L_i,x) : 1 \leq i \leq d\} \leq \min\{\tau(R_i,x) : 1 \leq i \leq d\} \tag{5.3}$$

and

$$\mathrm{M}\left(T_0 - \{w,v,v'\},x\right) \geq \mathrm{M}\left(T_0 - \{w',w,v\},x\right). \tag{5.4}$$

Proof. Let T be as in the lemma. By repeated application of (3.17) and (3.18) we obtain

$$\mathrm{M}(T,x) = G_{T_0}(L_1,\ldots,L_d;R_1,\ldots,R_d)\prod_{i=1}^{d}\mathrm{M}(L_i,x)\prod_{i=1}^{d}\mathrm{M}(R_i,x),$$

where

$$\begin{aligned}
&G_{T_0}(L_1,\ldots,L_d;R_1,\ldots,R_d)\\
&:= \left(1+x\sum_{i=1}^{d}\tau(L_i,x)\right)\left(1+x\sum_{i=1}^{d}\tau(R_i,x)\right)\mathrm{M}\left(T_0-\{w,v\},x\right)\\
&+\left(1+x\sum_{i=1}^{d}\tau(L_i,x)\right)x\,\mathrm{M}\left(T_0-\{w,v,v'\},x\right)\\
&+\left(1+x\sum_{i=1}^{d}\tau(R_i,x)\right)x\,\mathrm{M}\left(T_0-\{w',w,v\},x\right)+x^2\,\mathrm{M}\left(T_0-\{w',w,v,v'\},x\right).
\end{aligned}$$

Suppose that the $L_1, \ldots, L_d, R_1, \ldots, R_d$ do not satisfy (5.3), and we have the multiset equality

$$\{L_1, \ldots, L_d, R_1, \ldots, R_d\} = \{L_1', \ldots, L_d', R_1', \ldots, R_d'\}$$

such that $\tau(L_1', x) \le \cdots \le \tau(L_d', x) \le \tau(R_1', x) \le \cdots \le \tau(R_d', x)$. Then we have

$$G_{T_0}(L_1, \ldots, L_d; R_1, \ldots, R_d) > G_{T_0}(L_1', \ldots, L_d'; R_1', \ldots, R_d')$$

or

$$G_{T_0}(L_1, \ldots, L_d, R_1, \ldots, R_d) > G_{T_0}(R_1', \ldots, R_d', L_1', \ldots, L_d')$$

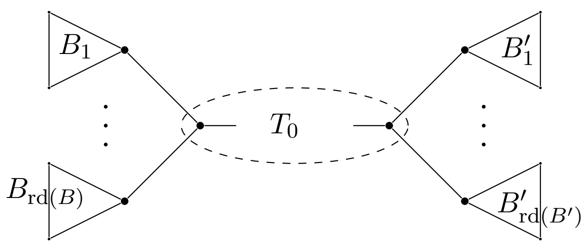
depending on whether $\mathrm{M}\left(T_0 - \{w, v, v'\}, x\right) \le \mathrm{M}\left(T_0 - \{w', w, v\}, x\right)$ or $\mathrm{M}\left(T_0 - \{w, v, v'\}, x\right) > \mathrm{M}\left(T_0 - \{w', w, v\}, x\right)$. This implies that if T' is the tree obtained from T by replacing L_i and R_i by L_i' and R_i' respectively for $i = 1, \ldots, d$, and T'' is obtained from T' by swapping L_i and R_i for all $i = 1, \ldots, d$, then we have $\mathrm{M}(T, x) > \mathrm{M}(T', x)$ or $\mathrm{M}(T, x) > \mathrm{M}(T'', x)$. This contradicts the minimality of $\mathrm{M}(T, x)$. Hence we can conclude that (5.3) must hold, which means that $T = T'$. Furthermore (5.4) must also hold, otherwise we would have

$$\begin{aligned}
&\mathrm{M}(T, x) - \mathrm{M}(T'', x) = \mathrm{M}(T', x) - \mathrm{M}(T'', x) \\
&= \left(\sum_{i=1}^{d} \tau(R_i, x) - \tau(L_i, x) \right) x^2 (\mathrm{M}(T_0 - \{w', w, v\}, x) - \mathrm{M}(T_0 - \{w, v, v'\}, x)) \\
&> 0,
\end{aligned}$$

which again contradicts the minimality of $\mathrm{M}(T, x)$. □

Taking into account the constraint of preserving the degree sequence, we restate the exchange lemma in a slightly different way. The proof is completely analogous.

Lemma 5.9 *Let $x > 0$ and let T be an x-minimal tree. Let $B = [B_1, \ldots, B_{\mathrm{rd}(B)}]$ and $B' = [B_1', \ldots, B_{\mathrm{rd}(B')}']$ be disjoint complete branches of T: the tree T can be decomposed as*



for some T_0. If

$$\min\{\tau(B_i, x) : 1 \le i \le \mathrm{rd}(B)\} < \max\{\tau(B_i', x) : 1 \le i \le \mathrm{rd}(B')\},$$

then we must have

$$\max\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} \le \min\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\},$$
$$\mathrm{M}(T_0 - \{\mathrm{r}(B),\mathrm{r}(B'),\mathrm{rn}(B')\}) \ge \mathrm{M}(T_0 - \{\mathrm{rn}(B),\mathrm{r}(B),\mathrm{r}(B')\})$$

and $\mathrm{rd}(B) \le \mathrm{rd}(B')$.

In Lemma 5.9 we have $\tau(B,x) > \tau(B',x)$ and $\mathrm{rd}(B) \le \mathrm{rd}(B')$, it turns out that this is not just a coincidence:

Lemma 5.10 *Let* $B = [B_1,\dots,B_{\mathrm{rd}(B)}]$ *and* $B' = [B'_1,\dots,B'_{\mathrm{rd}(B')}]$ *be two disjoint complete branches of an* x*-minimal tree for some* $x > 0$*. If* $\tau(B,x) > \tau(B',x)$*, then the inequalities*

$$\max\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} \le \min\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\}$$

and $\mathrm{rd}(B) \le \mathrm{rd}(B')$ *hold.*

Proof. Assume that

$$\tau(B,x) = \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B)} \tau(B_i,x)} > \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B')} \tau(B'_i,x)} = \tau(B',x). \tag{5.5}$$

Then it is impossible to have

$$\max\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} > \min\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\},$$

because using Lemma 5.9 it would lead to

$$\min\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} \ge \max\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\}$$

and hence $\tau(B,x) \le \tau(B',x)$, which is a contradiction with (5.5). Thus, we obtain

$$\max\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} \le \min\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\} \tag{5.6}$$

which implies

$$\min\{\tau(B_i,x) : 1 \le i \le \mathrm{rd}(B)\} \le \max\{\tau(B'_i,x) : 1 \le i \le \mathrm{rd}(B')\}. \tag{5.7}$$

If (5.7) was an equality, then combined with (5.6) it would imply $\tau(B_1,x) = \cdots = \tau(B_{\mathrm{rd}(B)},x) = \tau(B'_1,x) = \cdots = \tau(B'_{\mathrm{rd}(B')},x)$, therefore we must have $\mathrm{rd}(B) < \mathrm{rd}(B')$ for (5.5) to hold. Otherwise, (5.7) is a strict inequality and we use Lemma 5.9 to obtain $\mathrm{rd}(B) \le \mathrm{rd}(B')$. □

From the above lemmas, we know that the x-minimal trees are of the following type:

Definition 5.11 Let $x > 0$. We say that a tree T is *exchange-x-minimal* if and only if whenever $B = [B_1, \ldots, B_{\mathrm{rd}(B)}]$ and $B' = [B'_1, \ldots, B'_{\mathrm{rd}(B')}]$ are two disjoint complete branches in T, then one of the following must hold:

$i)$ $\mathrm{rd}(B) \leq \mathrm{rd}(B')$ and

$$\max\{\tau(B_1, x), \ldots, \tau(B_{\mathrm{rd}(B)}, x)\} \leq \min\{\tau(B'_1, x), \ldots, \tau(B'_{\mathrm{rd}(B')}, x)\},$$

$ii)$ $\mathrm{rd}(B) \geq \mathrm{rd}(B')$ and

$$\min\{\tau(B_1, x), \ldots, \tau(B_{\mathrm{rd}(B)}, x)\} \geq \max\{\tau(B'_1, x), \ldots, \tau(B'_{\mathrm{rd}(B')}, x)\}.$$

It follows directly from the definition that any exchange-x-minimal tree satisfies the two Lemmas 5.9 and 5.10, and each x-minimal tree is exchange-x-minimal.

Now, we can determine the degrees of all the vertices adjacent to a leaf in an exchange-x-minimal tree. Recall that $[d]$ denotes a pseudo-leaf branch and $\mathcal{P}(T)$ the set of all pseudo-leaves of T (see Definition 5.4).

Lemma 5.12 *Let $x > 0$ and let T be an exchange-x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$. Let S and S' be the subsets of $V(T)$ defined as follows: $v \in S$ if and only if v is adjacent to a leaf but v is not a pseudo-leaf, and $v \in S'$ if and only if v is adjacent to the root of a complete branch $[[d_1], \ldots, [d_{d_n-1}]]$. Then the following hold:*

$i)$ *If $v \in \mathcal{P}(T)$ and $w \in V(T) - \mathcal{P}(T)$, then we have $\deg(v) \geq \deg(w)$.*

$ii)$ *S contains at most one element.*

$iii)$ *If $v \in S$ and if $w \in V(T) - (S \cup \mathcal{P}(T))$, then we have $\deg(v) \geq \deg(w)$.*

$iv)$ *If $n \geq d_n + 1$, then S' is not empty and $\mathcal{P}(T)$ has at least d_n elements.*

$v)$ *If $n \geq d_n + 1$, then $S \subseteq S'$.*

Proof. Let $(d_1, \ldots, d_n)$ be the reduced degree sequence of T and assume that T is exchange-x-minimal for some $x > 0$. For a vertex $w \in V(T) - \mathcal{P}(T)$ which is not a leaf and for any vertex $w' \neq w$ in T, we define

$$B(w, w') \\ = \{B : B \text{ is a complete branch in } T, \mathrm{rn}(B) = w, \mathrm{rd}(B) \geq 1, w' \notin V(B)\}.$$

Since w is not a leaf nor a pseudo-leaf, it has at least two non-leaf branches, at least one of them does not contain w'. This shows that $B(w, w') \neq \varnothing$.

Assume that $v \in \mathcal{P}(T)$ and $w \in V(T) - \mathcal{P}(T)$. If w is a leaf, then obviously $\deg(v) \geq \deg(w) = 1$. Otherwise, we consider a complete branch $B \in B(w, v)$ and a leaf L adjacent to v. Then by Remark 5.7 we have $\tau(L, x) > \tau(B, x)$. By Lemma 5.9 this implies $\deg(v) \geq \deg(w)$. This shows $i)$.

Now, assume that a non-pseudo-leaf vertex w is adjacent to a leaf L. Let $u \in V(T)$ be a vertex adjacent to a leaf L' and $u \neq w$. Let $B \in B(w, u)$. Since $\tau(L', x) > \tau(B, x)$, by Lemma 5.9 we deduce that for all complete branches C attached to u either $w \in V(C)$ or $\tau(C, x) \geq \tau(L, x) = 1$ (meaning that C is a leaf). Hence, u is a pseudo-leaf and we deduce $ii)$.

Let $v \in S$ be a non-pseudo-leaf vertex adjacent to a leaf L and $w \in V(T) - (S \cup \mathcal{P}(T))$. If w is a leaf, then we trivially have $\deg(v) \geq \deg(w)$. Otherwise, w is not a leaf and we can consider a complete branch $B \in B(w, v)$. Since B is not a leaf, we get $\tau(L, x) > \tau(B, x)$, then by Lemma 5.9 we deduce $\deg(v) \geq \deg(w)$, which proves $iii)$.

As we will see in the proof of the next lemma, $i)$ and $ii)$ are enough to obtain a clear description of the exchange-x-minimal tree T if $n \leq d_n$. For the rest of the proof we assume that $n \geq d_n + 1$ (≥ 3). If T has a vertex v' which is not a leaf and not adjacent to a leaf, then each of the $\deg(v')$ ($\geq d_n$) branches of v' contains a pseudo-leaf. On the other hand, if all vertices in T which are not leaves are adjacent to a leaf, then by $ii)$ we deduce that there are $n - 1 \geq d_n$ pseudo-leaves in T. In either case, together with $i)$, this implies that we can form a set $\mathbb{B} = \{B_1 \approx_r [d_1], \ldots, B_{d_n - 1} \approx_r [d_{d_n - 1}]\}$ of $d_n - 1$ complete branches of T. Furthermore, for any branch B of T which is not isomorphic to an element of $\mathbb{B}$, the inequality

$$\tau(B_1, x) \leq \cdots \leq \tau(B_{d_n - 1}, x) < \tau(B, x) \tag{5.8}$$

holds, see Remark 5.7. Let S_1 be the set which contains all non-leaf vertices of T adjacent to a pseudo-leaf of degree d_1. Let S_2 be the subset of S_1 which contains all elements of S_1 with smallest degree. In view of (5.8) and $i)$, the choice of the elements of S_1 and Lemma 5.9 imply that for any $v \in S_2$ we have $\deg(v) = d_n$. Let $v_2 \in S_2$ be chosen to be adjacent to as many pseudo-leaves of degree d_1 as possible, and let $C_1, \ldots, C_{d_n}$ be the connected components of $T - v_2$ such that

$$\tau(C_1, x) \leq \cdots \leq \tau(C_{d_n}, x).$$

Note that $C_1 \approx_r B_1$. If $C_j \approx_r B_j$ for all $j \leq d_n - 1$, then $v_2 \in S'$. Otherwise, there exists $2 \leq i \leq d_n - 1$ for which $B_i \not\approx_r C_i$. Assume that

i is the smallest such integer. Hence we have $\tau(B_i, x) < \tau(C_i, x) \leq \cdots \leq \tau(C_{d_n}, x)$, which means that $B_i \not\approx_r C_j$ for all $j \geq i$. Hence, there must be a complete branch $B' \approx_r B_i$ of T such that $\mathrm{rn}(B') \neq v_2$. The case $\mathrm{rn}(B') \in \mathcal{P}(T)$ is impossible, it would imply $n = 2$. Hence $\mathrm{rn}(B')$ can only be contained in exactly one of $C_i, \ldots, C_{d_n}$. Let $C'_2 = B', C'_3, \ldots, C'_{\deg(\mathrm{rn}(B'))}$ be the connected components of $T - \mathrm{rn}(B')$ which do not contain v_2. By Lemma 5.9 the relation

$$\tau(B', x) = \tau(B_i, x) < \min\{\tau(C_{d_n-1}, x), \tau(C_{d_n}, x)\} \tag{5.9}$$

implies $\deg(v_2) \geq \deg(\mathrm{rn}(B'))$, thus $\deg(\mathrm{rn}(B')) = d_n$, and $\tau(C'_k, x) \leq \tau(C_1, x) = \tau(B_1, x)$ for all $k \geq 2$. In view of (5.8) this allows us to deduce that $C'_2 \approx_r \cdots \approx_r C'_{\deg(\mathrm{rn}(B'))} \approx_r B_1$. Hence $\mathrm{rn}(B') \in S'$. This proves that a complete branch $R_{d_n} \approx_r [[d_1], \ldots, [d_{d_n-1}]]$ always exists in T. With the condition $n \geq d_n + 1$, the tree T must have a non-leaf vertex in $T - R_{d_n}$, therefore at least one pseudo-leaf of T is in $T - R_{d_n}$. Hence, we have $iv)$.

Notice that for any non-leaf complete branch B of T not isomorphic to R_{d_n} we have

$$\tau(R_{d_n}, x) > \tau(B, x) \tag{5.10}$$

by (5.8). If S is empty, then $v)$ trivially holds. Now, we assume that $z \in S$. By contradiction, assume that none of the connected components of $T - z$ is isomorphic to R_{d_n}. Let L be a leaf attached to z and $B \in B(z, \mathrm{rn}(R_{d_n}))$. The inequality $\tau(L, x) > \tau(R_{d_n}, x) > \tau(B, x)$ contradicts Lemma 5.9. This completes the proof of $v)$. □

A rough picture of exchange-x-minimal trees can already be drawn from the information in Lemma 5.12.

Lemma 5.13 *Let T be an exchange-x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$. If $n \leq d_n + 1$ then T can be obtained by merging the root of each of the complete branches $[d_1], \ldots, [d_{n-1}]$ to a leaf of $[d_n + 1]$, see Figure 5.3 (a). Otherwise, T has the form shown in Figure 5.3 (b) for some integer $k \geq d_n$, where labels indicate degrees of the vertices and T_0 is a subgraph of T which does not contain any leaves or pseudo-leaves of T.*

Proof. If $n \leq d_n$, then it is impossible to have a non-leaf vertex of T not adjacent to a leaf. Hence, by $ii)$ of Lemma 5.12, at most one non-leaf vertex of T is allowed not to be a pseudo-leaf. It follows from $i)$ (of Lemma 5.12) that T is as in Figure 5.3 (a).

For $n = d_n + 1$, by $iv)$ of Lemma 5.12 we know that there is a vertex v in T adjacent to the root of the complete branch $[[d_1], \ldots, [d_{d_n-1}]]$. Consequently, v has to be a pseudo-leaf of degree d_{d_n}. This again means that T has the form in 5.3 (a). In fact, in this case, T also has

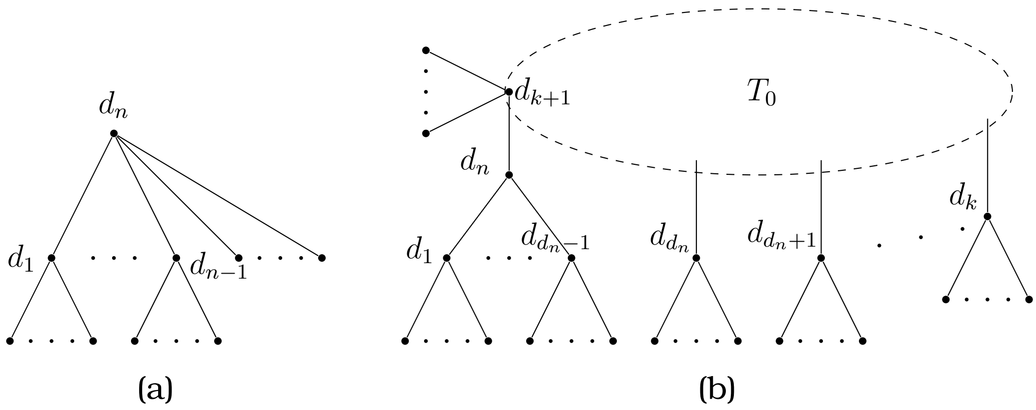


Figure 5.3: Tree T described in Lemma 5.13

the form in Figure 5.3 (b), but T_0 is reduced to a single vertex which has to be a pseudo-leaf.

For the case of $n \geq d_n+2$, Figure 5.3 (b) is just an easy translation of Lemma 5.12, where k is the number of pseudo-leaves in T. $\square$

From now on, for all trees T with reduced degree sequence $(d_1, \ldots, d_n)$, whenever d_i is used as label in a drawing of T, then d_i is the degree of the vertex to which it is assigned.

Generalizing (5.8), we will see in the next lemma that a complete branch B in an exchange-x-minimal tree can be identified, up to root-preserving isomorphism, by evaluating $\tau(B, x)$.

Lemma 5.14 *Let B and B' be two disjoint complete branches of an exchange-x-minimal tree for some $x > 0$. Then*

$$\tau(B, x) = \tau(B', x) \tag{5.11}$$

only if $B \approx_r B'$.

Proof. Without loss of generality we can assume that $\mathrm{h}(B') \leq \mathrm{h}(B)$. We reason by induction with respect to $\mathrm{h}(B)$. For $0 \leq \mathrm{h}(B) \leq 1$:

- If B' is a leaf then by Remark 5.7 the equality (5.11) can hold only if B is also a leaf.

- If B' is a pseudo-leaf branch, then so is B. In view of the formula $\tau([d+1], x) = 1/(xd+1)$, the equality (5.11) is possible only if $\mathrm{rd}(B') = \mathrm{rd}(B)$. This implies $B' \approx_r B$.

Assume that the lemma holds whenever $\mathrm{h}(B) \leq k$ for some positive integer k. Now, assume that $\mathrm{h}(B) = k+1$ and (5.11) holds. Since $\mathrm{h}(B) \geq 2$, it follows that $\mathrm{r}(B)$ cannot be a pseudo-leaf, otherwise B' would have to be a leaf by disjointness (the unique non-leaf vertex adjacent to $\mathrm{r}(B)$, if there is such a vertex, would be in B) and contradict (5.11), see Remark 5.7 again. If B' is a pseudo-leaf branch, then $\mathrm{rd}(B') \geq \mathrm{rd}(B)$ by Lemma 5.12 and thus

$$\tau(B,x) > \frac{1}{x\,\mathrm{rd}(B)+1} \geq \frac{1}{x\,\mathrm{rd}(B')+1} = \tau(B',x).$$

For the rest of this proof we assume that $\mathrm{h}(B') \geq 2$. Write $B = [B_1, \dots, B_{\mathrm{rd}(B)}]$ and $B' = [B'_1, \dots, B'_{\mathrm{rd}(B')}]$. Then we have

$$\tau(B,x) = \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B)} \tau(B_i,x)} \text{ and } \tau(B',x) = \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B')} \tau(B'_i,x)}.$$

Hence, (5.11) implies

$$\sum_{i=1}^{\mathrm{rd}(B')} \tau(B_i,x) = \sum_{i=1}^{\mathrm{rd}(B)} \tau(B'_i,x). \tag{5.12}$$

If $\tau(B_1,x) = \cdots = \tau(B_{\mathrm{rd}(B)},x) = \tau(B'_1,x) = \cdots = \tau(B'_{\mathrm{rd}(B')})$, then for (5.12) to hold we must also have $\mathrm{rd}(B) = \mathrm{rd}(B')$, since $\tau(B_i,x) > 0$ and $\tau(B'_j,x) > 0$ for all i and j. By the induction hypothesis there are isomorphisms $f_i : B_i \to B'_i$, $1 \leq i \leq \mathrm{rd}(B)$, which preserve roots. The function f defined by

$$\begin{aligned} f : V(B) &\longrightarrow V(B') \\ x &\longmapsto \begin{cases} f_i(x) \text{ if } x \in V(B_i) \\ \mathrm{r}(B') \text{ if } x = \mathrm{r}(B) \end{cases} \end{aligned}$$

is clearly an isomorphism, it preserves roots by construction.

Otherwise, there exists $i_0 \in \{1, 2, \dots, \mathrm{rd}(B)\}$ and $j_0 \in \{1, 2, \dots, \mathrm{rd}(B')\}$ such that $\tau(B_{i_0},x) < \tau(B'_{j_0},x)$ or $\tau(B_{i_0},x) > \tau(B'_{j_0},x)$. Assume that

$$\tau(B_{i_0},x) < \tau(B'_{j_0},x). \tag{5.13}$$

By Lemma 5.9 this implies that

$$\max\{\tau(B_i,x) : 1 \leq i \leq \mathrm{rd}(B)\} \leq \min\{\tau(B'_i,x) : 1 \leq i \leq \mathrm{rd}(B')\} \tag{5.14}$$

and

$$\mathrm{rd}(B) \leq \mathrm{rd}(B'). \tag{5.15}$$

But (5.13), (5.14) and (5.15) lead to a contradiction with (5.12). Similar reasoning can be used to show that $\tau(B_{i_0},x) > \tau(B'_{j_0},x)$ also leads to a contradiction with (5.12). □

5.2 Uniqueness of the x-minimal tree

In this section we will prove that for all $x > 0$, any two exchange-x-minimal trees with the same degree sequence are isomorphic. This implies that there is a unique, up to isomorphism, x-minimal tree for any given degree sequence.

Let T be a tree. For any two elements v and v' of $\mathcal{P}(T)$, there exists a unique path which joins v and v' in T. Given a non-negative integer $j < \mathrm{d}(v, v')$, we denote by $B_{vv'}(j)$ the unique complete branch of T such that $v \in V(B_{vv'}(j))$, $v' \notin V(B_{vv'}(j))$ and $\mathrm{d}(v, \mathrm{r}(B_{vv'}(j))) = j$, see Figure 5.4 for some examples.

We write $v \prec_{k,x} v'$ if and only if we have $B_{vv'}(k) \cap B_{v'v}(k) = \varnothing$, and in addition, k is even and $\tau(B_{vv'}(k), x) < \tau(B_{v'v}(k), x)$ or k is odd and $\tau(B_{vv'}(k), x) > \tau(B_{v'v}(k), x)$. Otherwise we write $v \not\prec_{k,x} v'$ (which simply means "not $v \prec_{k,x} v'$"). Immediate properties of $\prec_{k,x}$ and $\not\prec_{k,x}$ are as follows:

Lemma 5.15 *Let $x > 0$, and let T be an exchange-x-minimal tree. Let $v_1, \ldots, v_m$ be m (≥ 2) elements of $\mathcal{P}(T)$, and let k be an integer such that*

i) $B_{v_1v_2}(k) \cap B_{v_2v_1}(k) = \cdots = B_{v_{m-1}v_m}(k) \cap B_{v_mv_{m-1}}(k) = \varnothing$,

ii) $B_{v_1v_2}(k) \approx_r B_{v_2v_1}(k), \ldots, B_{v_{m-1}v_m}(k) \approx_r B_{v_mv_{m-1}}(k)$.

Then it is impossible to have

$$v_1 \prec_{k+1,x} \cdots \prec_{k+1,x} v_m \prec_{k+1,x} v_1. \tag{5.16}$$

Proof. For the case where $m = 2$, it is impossible to have

$$v_1 \prec_{k+1,x} v_2 \prec_{k+1,x} v_1,$$

because it leads to the absurd inequalities

$$\tau(B_{v_1v_2}(k+1), x) < \tau(B_{v_2v_1}(k+1), x) < \tau(B_{v_1v_2}(k+1), x).$$

For the other values of m we reason by induction. The case of $m = 3$ will play a key role, so we prove it first. Assume that v_1, v_2, v_3 are three elements of $\mathcal{P}(T)$ such that $i)$ and $ii)$ hold and

$$v_1 \prec_{k+1,x} v_2 \prec_{k+1,x} v_3. \tag{5.17}$$

We will show that $v_1 \prec_{k+1,x} v_3$ or $B_{v_1v_3}(k+1) \approx_r B_{v_3v_1}(k+1)$, both cases are in contradiction with $v_3 \prec_{k+1,x} v_1$.

T can be decomposed as in Figure 5.4 for some tree T_0: there exists a subgraph of T which is a tree with exactly three leaves, and they are v_1, v_2 and v_3. We must have $k \leq \min\{i, j, l\}$. This is because of the following:

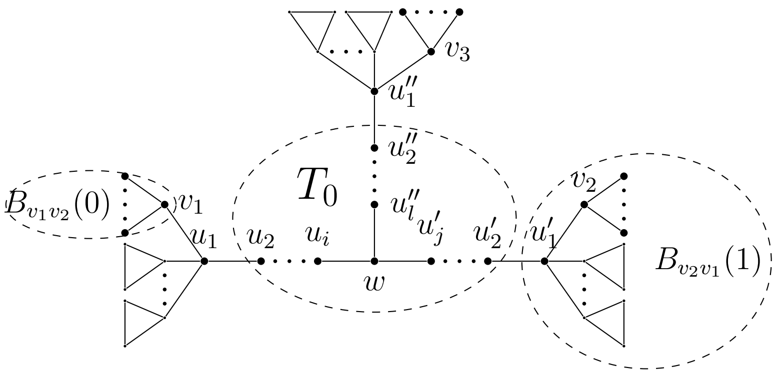


Figure 5.4: Decomposition of T for the proof of Lemma 5.15

- If $k \geq i+1$ (resp. $k \geq l+1$), then for the part $v_1 \prec_{k+1,x} v_2$ (resp. $v_2 \prec_{k+1,x} v_3$) of (5.17) to hold we must also have $k < j$. This implies $B_{v_2v_3}(k) = B_{v_2v_1}(k)$ and
$$B_{v_2v_3}(k+1) = B_{v_2v_1}(k+1). \tag{5.18}$$
Hence, we have
$$B_{v_3v_2}(k) \approx_r B_{v_2v_3}(k) = B_{v_2v_1}(k) \approx_r B_{v_1v_2}(k). \tag{5.19}$$
If $i > l$, then $B_{v_1v_2}(k)$ is a proper subgraph of $B_{v_3v_2}(k)$. Similarly, if $l > i$, then $B_{v_3v_2}(k)$ is a proper subgraph of $B_{v_1v_2}(k)$. In both cases (5.19) is absurd. For $i = l$ we have $B_{v_1v_2}(k+1) = B_{v_3v_2}(k+1)$, with (5.18) this contradicts (5.17).

- If $k \geq j+1$, then for (5.17) to hold, we must have $k < i$ and $k < l$. In this case, $B_{v_3v_2}(k)$ and $B_{v_1v_2}(k)$ are proper subgraphs of $B_{v_2v_1}(k)$ and $B_{v_2v_3}(k)$, respectively. Therefore, the two relations $B_{v_3v_2}(k) \approx_r B_{v_2v_3}(k)$ and $B_{v_1v_2}(k) \approx_r B_{v_2v_1}(k)$ obtained from $ii)$ would not hold simultaneously.

Since $k \leq \min\{i, j, l\}$, we know that $\mathrm{rd}(B_{v_1v_2}(k+1)) = \mathrm{rd}(B_{v_1v_3}(k+1))$, $\mathrm{rd}(B_{v_2v_1}(k+1)) = \mathrm{rd}(B_{v_2v_3}(k+1))$, $\mathrm{rd}(B_{v_3v_1}(k+1)) = \mathrm{rd}(B_{v_3v_2}(k+1))$, and $ii)$ implies that
$$B_{v_1v_3}(k) \approx_r B_{v_3v_1}(k). \tag{5.20}$$
By application of Lemma 5.10, we can deduce from (5.17) that
$$\mathrm{rd}(B_{v_1v_2}(k+1)) \geq \mathrm{rd}(B_{v_2v_1}(k+1)) = \mathrm{rd}(B_{v_2v_3}(k+1)) \geq \mathrm{rd}(B_{v_3v_2}(k+1))$$
if $k+1$ is even and
$$\mathrm{rd}(B_{v_1v_2}(k+1)) \leq \mathrm{rd}(B_{v_2v_1}(k+1)) = \mathrm{rd}(B_{v_2v_3}(k+1)) \leq \mathrm{rd}(B_{v_3v_2}(k+1))$$

if $k+1$ is odd. In each case, if at least one of the two inequalities is strict then we get

$$\mathrm{rd}(B_{v_1v_3}(k+1)) = \mathrm{rd}(B_{v_1v_2}(k+1)) > \mathrm{rd}(B_{v_3v_2}(k+1)) = \mathrm{rd}(B_{v_3v_1}(k+1))$$

if $k+1$ is even, and for the case of odd $k+1$ we have

$$\mathrm{rd}(B_{v_1v_3}(k+1)) = \mathrm{rd}(B_{v_1v_2}(k+1)) < \mathrm{rd}(B_{v_3v_2}(k+1)) = \mathrm{rd}(B_{v_3v_1}(k+1)).$$

In view of Lemmas 5.10 and 5.14, these can only lead to

$$\tau(B_{v_1v_3}(k+1), x) < \tau(B_{v_3v_1}(k+1), x)$$

for even $k+1$, and

$$\tau(B_{v_1v_3}(k+1), x) > \tau(B_{v_3v_1}(k+1), x)$$

for odd $k+1$. Therefore $v_1 \prec_{k+1,x} v_3$, but not $v_3 \prec_{k+1,x} v_1$. Now we are left with the situation where

$$\begin{aligned} d := \mathrm{rd}(B_{v_1v_2}(k+1)) &= \mathrm{rd}(B_{v_2v_1}(k+1)) \\ &= \mathrm{rd}(B_{v_2v_3}(k+1)) = \mathrm{rd}(B_{v_3v_2}(k+1)). \end{aligned} \tag{5.21}$$

The two cases depending on the parity of $k+1$ have to be treated separately:
Case 1: Assume that $k+1$ is even, so that (5.17) implies

$$\tau(B_{v_1v_2}(k+1), x) < \tau(B_{v_2v_1}(k+1), x) \tag{5.22}$$

and

$$\tau(B_{v_2v_3}(k+1), x) < \tau(B_{v_3v_2}(k+1), x). \tag{5.23}$$

If $k < \min\{i, j, l\}$, then $B_{v_1v_2}(k+1) = B_{v_1v_3}(k+1)$, $B_{v_2v_3}(k+1) = B_{v_2v_1}(k+1)$ and $B_{v_3v_2}(k+1) = B_{v_3v_1}(k+1)$. Hence, $v_1 \prec_{k+1} v_2 \prec_{k+1} v_3$ implies

$$\begin{aligned} \tau(B_{v_1v_3}(k+1), x) = \tau(B_{v_1v_2}(k+1), x) &< \tau(B_{v_2v_1}(k+1), x) \\ = \tau(B_{v_2v_3}(k+1), x) &< \tau(B_{v_3v_2}(k+1), x) = \tau(B_{v_3v_1}(k+1), x). \end{aligned}$$

With the obvious relation $B_{v_1v_3}(k+1) \cap B_{v_3v_1}(k+1) = \varnothing$, this leads to $v_1 \prec_{k+1,x} v_3$.

Otherwise, we have $k = \min\{i, j, l\}$. Let $B_{v_1v_2}(k+1) := [B_{v_1v_3}(k) = B_{v_1v_2}(k) = A_1, \ldots, A_d]$ and $B_{v_3v_2}(k+1) := [B_{v_3v_2}(k) = B_{v_3v_1}(k) = C_1, \ldots, C_d]$. There are two sub-cases:

- If $k = i = \min\{i,j,l\}$ or $k = l = \min\{i,j,l\}$, then for (5.17) to be possible we must have $k < j$. It follows that

$$B_{v_2v_1}(k+1) = B_{v_2v_3}(k+1) := [B_{v_2v_1}(k) = B_{v_2v_3}(k) = B_1, B_2, \dots, B_d]$$

and using $ii)$ we get

$$B_{v_1v_2}(k) \approx_r B_{v_2v_1}(k) = B_{v_2v_3}(k) \approx_r B_{v_3v_2}(k). \tag{5.24}$$

From (5.21), (5.22), (5.23) and Lemma 5.10 we obtain

$$\begin{aligned}&\max\{\tau(A_t,x) : d \geq t \geq 1\}\\ &\qquad > \min\{\tau(B_t,x) : d \geq t \geq 1\} \geq \max\{\tau(C_t,x) : d \geq t \geq 1\}\end{aligned}$$

and

$$\begin{aligned}&\min\{\tau(A_t,x) : d \geq t \geq 1\}\\ &\qquad \geq \max\{\tau(B_t,x) : d \geq t \geq 1\} > \min\{\tau(C_t,x) : d \geq t \geq 1\},\end{aligned}$$

which imply

$$\max\{\tau(A_t,x) : d \geq t \geq 1\} > \tau(B_{v_3v_1}(k),x) = \tau(C_1,x) \tag{5.25}$$

and

$$\tau(B_{v_1v_3}(k),x) = \tau(A_1,x) > \min\{\tau(C_t,x) : d \geq t \geq 1\}, \tag{5.26}$$

respectively. The situation $i = l = k$ is impossible, because it would lead to $B_{v_3v_2}(k+1) = B_{v_1v_2}(k+1)$, which is in contradiction with (5.17). Hence $k < \max\{i,l\}$. If $k < i$, then $B_{v_1v_2}(k+1) = B_{v_1v_3}(k+1)$ and (5.25) implies $v_1 \prec_{k+1,x} v_3$. If $k < l$, then $B_{v_3v_2}(k+1) = B_{v_3v_1}(k+1)$ and (5.26) implies $v_1 \prec_{k+1,x} v_3$.

- If $k = j = \min\{i,j,l\}$, then necessarily $k < i$ and $k < l$ for (5.17) to be possible. Note that in this case we get

$$B_{v_1v_3}(k+1) = B_{v_1v_2}(k+1), B_{v_3v_1}(k+1) = B_{v_3v_2}(k+1),$$

$B_{v_2v_3}(k) = B_{v_2v_1}(k)$, and hence by Lemma 5.10 the relations (5.22) and (5.23) imply

$$\begin{aligned}\max\{\tau(C_s,x) : d \geq s \geq 1\} &\leq \tau(B_{v_2v_3}(k),x) = \tau(B_{v_2v_1}(k),x)\\ &\leq \min\{\tau(A_s,x) : d \geq s \geq 1\}.\end{aligned} \tag{5.27}$$

If there exists t such that $\tau(A_t,x) > \tau(B_{v_3v_1}(k),x)$ or $\tau(B_{v_1v_3}(k),x) > \tau(C_t,x)$, then Lemma 5.9 allows us to deduce that

$$\tau(B_{v_1v_3}(k+1),x) = \tau(B_{v_1v_2}(k+1),x) < \tau(B_{v_3v_1}(k+1),x)$$

or

$$\tau(B_{v_1v_3}(k+1),x) < \tau(B_{v_3v_2}(k+1),1) = \tau(B_{v_3v_1}(k+1),x),$$

and this implies $v_1 \prec_{k+1,x} v_3$. Otherwise, with use of (5.27), we have

$$\begin{aligned}\tau(B_{v_1v_3}(k),x) &\le \min\{\tau(C_s,x) : d \ge s \ge 1\}\\ &\le \max\{\tau(C_s,x) : d \ge s \ge 1\} \le \tau(B_{v_2v_1}(k),x)\\ &\le \min\{\tau(A_s,x) : d \ge s \ge 1\}\\ &\le \max\{\tau(A_s,x) : d \ge s \ge 1\} \le \tau(B_{v_3v_1}(k),x).\end{aligned}$$

Since $B_{v_1v_3}(k) \approx_r B_{v_2v_1}(k) \approx_r B_{v_3v_1}(k)$, it follows that

$$B_{v_2v_1}(k) = B_{v_2v_3}(k) \approx_r A_1 \approx_r \cdots \approx_r A_d \approx_r C_1 \approx_r \cdots \approx_r C_d \tag{5.28}$$

and consequently

$$B_{v_1v_3}(k+1) \approx_r B_{v_3v_1}(k+1), \tag{5.29}$$

therefore $v_3 \not\prec_{k+1,x} v_1$. This is the single case where the relation $v_1 \prec_{k+1,x} v_3$ may not hold, but what we have seen shows that $v_3 \prec_{k+1,x} v_1$ cannot happen either. We write $S_k(v_1,v_2,v_3)$ if and only if v_1, v_2, v_3 satisfy this situation: that is, T can be decomposed as in Figure 5.5 (a) where $v_1 \in V(A_1)$, $v_2 \in V(B_2)$, $v_3 \in V(C_1)$ and the A_i's and C_i's satisfy (5.28).

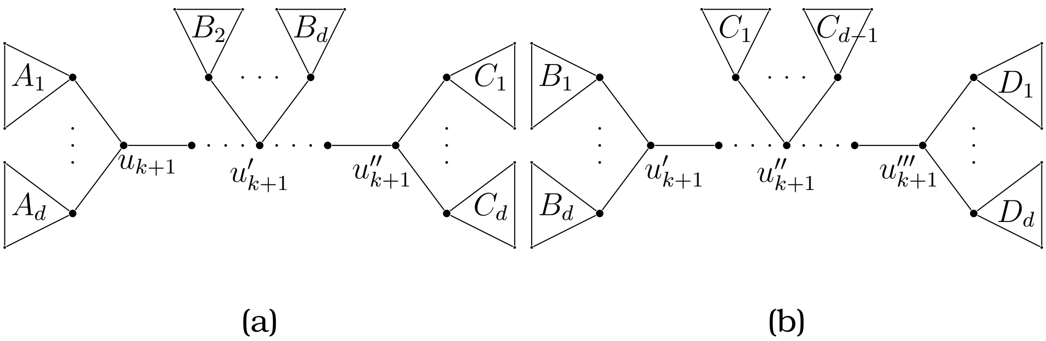


Figure 5.5: Two decompositions of T for the proof of Lemma 5.15

Case 2: Assume that $k+1$ is odd. Then we proceed exactly as in the previous case, but reverse any inequality involving τ, replace $\min$ and $\max$ by $\max$ and $\min$, respectively, except for $\min\{i,j,l\}$ and $\min\{i,l\}$.

We now assume that $m \ge 4$ and (5.16) holds. If $v_1 \prec_{k+1,x} v_3$ (resp. $v_2 \prec_{k+1,x} v_4$), we can remove v_2 (resp. v_3) from the list of v_i's to obtain a shorter list which satisfies i), ii) (see (5.20)) and (5.16). This would contradict the induction hypothesis. Hence, we must have $S_k(v_1, v_2, v_3)$ and $S_k(v_2, v_3, v_4)$, and then T has the two decompositions in Figure 5.5 by appropriately rearranging the indices, where $v_1 \in V(A_1)$, $v_2 \in V(B_2)$, $v_3 \in V(C_1)$, $v_4 \in V(D_1)$,

$$B_2 \approx_r A_1 \approx_r \cdots \approx_r A_d \approx_r C_1 \approx_r \cdots \approx_r C_d$$

and

$$C_1 \approx_r B_1 \approx_r \cdots \approx_r B_d \approx_r D_1 \approx_r \cdots \approx_r D_d.$$

But this is absurd because D_1 is a proper subgraph of C_d. □

For any $x > 0$ and for any exchange-x-minimal tree T with reduced degree sequence $(d_1, \ldots, d_n)$ we set $\mathcal{P}_0(T) = \mathcal{P}(T)$, and for any positive integer k we define

$$\mathcal{P}_k(T) = \mathcal{P}_{k-1}(T) - \{v \in \mathcal{P}_{k-1}(T) : v' \prec_{k-1,x} v \text{ for some } v' \in \mathcal{P}_{k-1}(T)\}.$$

In particular $\mathcal{P}_1(T)$ contains all the pseudo-leaves of degree d_1. The sequence of $\mathcal{P}_k(T)$ is stationary, for all k greater than $\lceil \mathrm{diam}(T)/2 \rceil$ we have $\mathcal{P}_k(T) = \mathcal{P}_{k+1}(T)$. In view of Lemmas 5.14 and 5.15 we know that in the process of passing from $\mathcal{P}_k(T)$ to $\mathcal{P}_{k+1}(T)$ it is impossible to have a cyclic chain of elimination, this implies that $\mathcal{P}_k(T) \neq \varnothing$ for all non-negative k. If $n \geq d_n + 1$, then any element left in $\mathcal{P}_{\lceil \mathrm{diam}(T)/2 \rceil}(T)$ is contained in a complete branch isomorphic to $R_{d_n}(T) = [[d_1], \ldots, [d_{d_n - 1}]]$ (see (5.10)), which is known to exist in T from the proof of Lemma 5.12. That is why we can define the following notations which will be used often in the rest of this chapter: $\nu_{d_1}(T)$ is an arbitrarily chosen element of $\mathcal{P}_{\lceil \mathrm{diam}(T)/2 \rceil}(T)$, $\nu_{d_n}(T)$ is the non-leaf vertex adjacent to $\nu_{d_1}(T)$, and $\nu(T)$ is the non-pseudo-leaf vertex adjacent to $\nu_{d_n}(T)$ (see Figure 5.6).

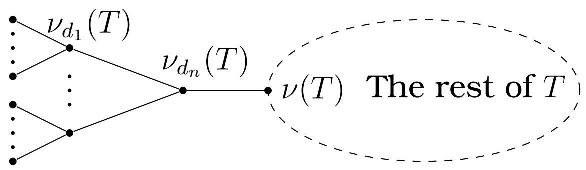


Figure 5.6: Place of $\nu_{d_1}(T), \nu_{d_n}(T)$ and $\nu(T)$ in T

Remark 5.16 Let T be an exchange-x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$, where $n \geq d_n + 1$. If $S \subseteq S'$ are the two sets as defined in Lemma 5.12, then clearly $\nu(T) \in S'$. If B and B' are disjoint complete branches such that $B \approx_r R_{d_n}(T) \approx_r B'$, $\mathrm{rn}(B) \in S$

and $\mathrm{rn}(B') \in S' - S$, then $\mathrm{rn}(B)$ has a leaf branch and hence knowing Lemma 5.9 we deduce that for any pseudo-leaves $v \in V(B)$, $v' \in V(B')$ and for any real $x > 0$ we have $\tau(B_{vv'}(2), x) < \tau(B_{v'v}(2), x)$ meaning that $v \prec_{2,x} v'$. This shows that $\nu(T) \in S$ if $S \neq \varnothing$.

Lemma 5.17 *For some $x > 0$, let T be an exchange-x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$. Let $B = [B_1, \ldots, B_{\mathrm{r}(B)}]$ and $B' = [B'_1, \ldots, B'_{\mathrm{r}(B')}]$ be two disjoint complete branches of T such that a pseudo-leaf in B_1 can be chosen to be $\nu_{d_1}(T)$ and $B_1 \approx_r B'_1$. Then the following hold:*

- *If $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B))$ is even, then $\mathrm{rd}(B) \geq \mathrm{rd}(B')$ and*

$$\min\{\tau(B_i, x) : 1 \leq i \leq \mathrm{rd}(B)\} \geq \max\{\tau(B'_i, x) : 1 \leq i \leq \mathrm{rd}(B')\}.$$

- *If $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B))$ is odd, then $\mathrm{rd}(B) \leq \mathrm{rd}(B')$ and*

$$\max\{\tau(B_i, x) : 1 \leq i \leq \mathrm{rd}(B)\} \leq \min\{\tau(B'_i, x) : 1 \leq i \leq \mathrm{rd}(B')\}.$$

Proof. Otherwise using Lemma 5.9 we would get $\tau(B', x) < \tau(B, x)$ for even $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B))$ and $\tau(B', x) > \tau(B, x)$ for odd $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B))$. This implies that $\nu_{d_1}(T) \prec_{\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B)), x} f(\nu_{d_1}(T))$, where $f : B_1 \to B'_1$ is an isomorphism preserving roots. This contradicts the choice of $\nu_{d_1}(T)$.

Note that, by Lemma 5.9 again, $B \approx_r B'$ can only happen if $\tau(B_1, x) = \cdots = \tau(B_{\mathrm{rd}(B)}, x) = \tau(B'_1) = \cdots = \tau(B'_{\mathrm{rd}(B)}, x)$. ☐

For any pseudo-leaf v of a tree T, we denote by $T - \langle v \rangle$ the tree obtained from T by removing the pseudo-leaf branch with root v. See Figure 5.7 for an example. We extend this notation to all subgraphs S of T, even if they do not contain v, in which case $S - \langle v \rangle = S$. The next lemma describes a specific proper subgraph of an exchange-x-minimal tree, which can inherit the property of exchange-x-minimality. This lemma plays an important role in inductive proofs of the main theorems of this chapter.

Lemma 5.18 *Let $x > 0$ and let T be an exchange-x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$. Then $T - \langle \nu_{d_1}(T) \rangle$ (see Figure 5.7) is an exchange-x-minimal tree.*

Note that $\nu(T)$ and $\nu(T - \langle \nu_{d_1}(T) \rangle)$ do not have to coincide. For instance, if $\nu(T)$ is adjacent to $\deg(\nu(T)) - 2$ leaves and $d_n = 2$, then $\nu(T)$ becomes a pseudo-leaf in $T - \langle \nu_{d_1}(T) \rangle$, while by definition $\nu(T - \langle \nu_{d_1}(T) \rangle)$ is not a pseudo-leaf.

Proof. $\nu(T)$ is the only non-pseudo-leaf vertex of T which can possibly become a pseudo-leaf in $T - \langle \nu_{d_1}(T) \rangle$. It is also the only vertex of $T - \langle \nu_{d_1}(T) \rangle$ which can possibly be a non-pseudo-leaf vertex adjacent

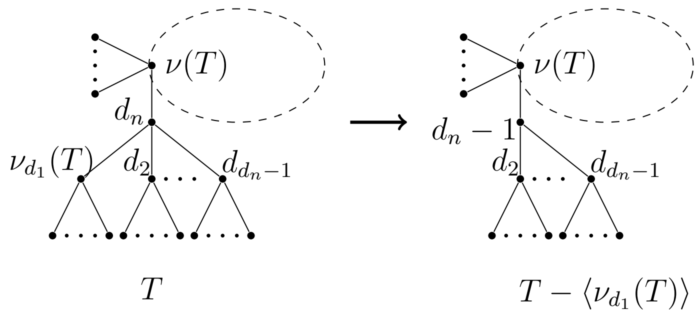


Figure 5.7: Construction of $T - \langle \nu_{d_1}(T) \rangle$

to a leaf. Therefore, since T satisfies Lemma 5.12, at least $i), ii)$ and $iii)$ of the same lemma are satisfied by $T - \langle \nu_{d_1}(T) \rangle$.

Furthermore, we claim the following for all complete branches $B_1 = [B_{1,1}, \ldots, B_{1,\mathrm{rd}(B_1)}]$ and $B_2 = [B_{2,1}, \ldots, B_{2,\mathrm{rd}(B_2)}]$ of T:

$i)$ If for some $i_0 \in \{1, 2\}$ we have $\nu_{d_1}(T) \in B_{i_0}$ and $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0}))$ is even, then $\tau(B_{i_0}, x) > \tau(B_{i_0} - \langle \nu_{d_1}(T) \rangle, x)$; and if $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0}))$ is odd, then $\tau(B_{i_0}, x) < \tau(B_{i_0} - \langle \nu_{d_1}(T) \rangle, x)$.

$ii)$ If $\tau(B_1, x) < \tau(B_2, x)$, then $\tau(B_1 - \langle \nu_{d_1}(T) \rangle, x) \leq \tau(B_2 - \langle \nu_{d_1}(T) \rangle, x)$.

Claim $i)$ can be shown by induction. Assume that $\nu_{d_1}(T) \in B_{i_0}$ for some $i_0 \in \{1, 2\}$. If $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0})) = 0$, i.e. $\nu_{d_1}(T) = \mathrm{r}(B_{i_0})$, then we have

$$\tau(B_{i_0}, x) > \tau(B_{i_0} - \langle \nu_{d_1}(T) \rangle, x) = \tau(\varnothing, x) = 0.$$

The induction step follows from relation (5.2).

For claim $ii)$ the case where B_1 and B_2 do not contain $\nu_{d_1}(T)$ is trivial. Assume that $i_0 \in \{1, 2\}$, $\nu_{d_1}(T) \in V(B_{i_0})$ and

$$\tau(B_1, x) < \tau(B_2, x). \tag{5.30}$$

By Lemmas 5.9 and 5.10 it follows from (5.30) that

$$\mathrm{rd}(B_1) \geq \mathrm{rd}(B_2) \tag{5.31}$$

and

$$\min\{\tau(B_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B_1)\} \geq \max\{\tau(B_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B_2)\}. \tag{5.32}$$

Consider the case where $\nu_{d_1}(T) = \mathrm{r}(B_{i_0})$. If $i_0 = 1$, then $B_1 - \langle \nu_{d_1}(T) \rangle$ is empty and $0 = \tau(B_1 - \langle \nu_{d_1}(T) \rangle, x) < \tau(B_2, x) = \tau(B_2 - \langle \nu_{d_1}(T) \rangle, x)$. It is impossible that $i_0 = 2$ in this case in view of (5.8). Assume that the claim holds for $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0})) = k$. Now, assume that $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0})) = k + 1$. We start with the case where k is even, i.e. $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{i_0}))$ is odd:

- If $i_0 = 1$, without loss of generality we can assume that $\nu_{d_1}(T) \in B_{1,1}$. There are two sub-cases:

  – If $\tau(B_{2,1}, x) < \tau(B_{1,1}, x)$, then by the induction hypothesis we have $\tau(B_{2,1}, x) = \tau(B_{2,1} - \langle \nu_{d_1}(T)\rangle, x) \leq \tau(B_{1,1} - \langle \nu_{d_1}(T)\rangle, x)$. Using (5.31) and (5.32) we deduce that

$$\begin{aligned}\tau(B_1 - \langle \nu_{d_1}(T)\rangle, x) &= \frac{1}{1 + x\tau(B_{1,1} - \langle \nu_{d_n}(T)\rangle, x) + x\sum_{i=2}^{\mathrm{rd}(B_1)} \tau(B_{1,i}, x)} \\ &\leq \frac{1}{1 + x\tau(B_{2,1} - \langle \nu_{d_n}(T)\rangle, x) + x\sum_{i=2}^{\mathrm{rd}(B_2)} \tau(B_{2,i}, x)} \\ &\leq \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B_2)} \tau(B_{2,i}, x)} \\ &= \tau(B_2, x) = \tau(B_2 - \langle \nu_{d_1}(T)\rangle, x).\end{aligned}$$

  – Otherwise $\tau(B_{2,1}, x) = \tau(B_{1,1}, x)$, which implies that $B_{2,1} \approx_r B_{1,1}$, see Lemma 5.14. According to Lemma 5.17, the following must hold:

$$\mathrm{rd}(B_1) \leq \mathrm{rd}(B_2) \tag{5.33}$$

  and

$$\begin{aligned}\max\{\tau(B_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B_1)\} \\ \leq \min\{\tau(B_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B_2)\}.\end{aligned} \tag{5.34}$$

  But (5.31), (5.32), (5.33) and (5.34) imply $\mathrm{rd}(B_1) = \mathrm{rd}(B_2)$,

$$\tau(B_{1,i}, x) = \cdots = \tau(B_{1,\mathrm{rd}(B_1)}, x) = \tau(B_{2,1}, x) = \cdots = \tau(B_{2,\mathrm{rd}(B_2)}, x)$$

  and contradict (5.30).

- For $i_0 = 2$ we can assume that $\nu_{d_1}(T) \in V(B_{2,1})$. Then by $i)$ we have $\tau(B_{2,1}, x) > \tau(B_{2,1} - \langle \nu_{d_1}(T)\rangle, x)$ and thus

$$\begin{aligned}\tau(B_2 - \langle \nu_{d_1}(T)\rangle, x) &= \frac{1}{1 + x\tau(B_{2,1} - \langle \nu_{d_1}(T)\rangle, x) + x\sum_{i=2}^{\mathrm{rd}(B_2)} \tau(B_{2,i}, x)} \\ &> \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B_2)} \tau(B_{2,i}, x)} \\ &= \tau(B_2, x) > \tau(B_1, x) = \tau(B_1 - \langle \nu_{d_1}(T)\rangle, x).\end{aligned}$$

The case of odd k can be treated in the same way, by reversing all inequalities and replacing $B_1, B_2, B_{1,i}, B_{2,i}, i_0 = 1, i_0 = 2$ with $B_2, B_1, B_{2,i}, B_{1,i}, i_0 = 2, i_0 = 1$ respectively for all i.

It is left to show that $i)$ and $ii)$ imply exchange-x-minimality of $T - \langle \nu_{d_1}(T)\rangle$. Let $B_1' = [B_{1,1}', \ldots, B_{1,\mathrm{rd}(B_1')}']$ and $B_2' = [B_{2,1}', \ldots, B_{2,\mathrm{rd}(B_2')}']$ be two disjoint complete non-leaf branches of $T - \langle \nu_{d_1}(T) \rangle$. Then there exist complete branches $B_1 = [B_{1,1}, \ldots, B_{1,\mathrm{rd}(B_1)}]$ and $B_2 = [B_{2,1}, \ldots, B_{2,\mathrm{rd}(B_2)}]$ of T such that $B'_i = B_i - \langle \nu_{d_1}(T) \rangle$, $i \in \{1, 2\}$. Without loss of generality, we can assume that

$$\mathrm{rd}(B_1) \leq \mathrm{rd}(B_2) \tag{5.35}$$

and

$$\max\{\tau(B_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B_1)\} \\ \leq \min\{\tau(B_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B_2)\}. \tag{5.36}$$

We aim to show that one of the following holds:

$a)$ $\mathrm{rd}(B'_1) \leq \mathrm{rd}(B'_2)$ and

$$\max\{\tau(B'_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_1)\} \leq \min\{\tau(B'_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_2)\},$$

$b)$ $\mathrm{rd}(B'_1) \geq \mathrm{rd}(B'_2)$ and

$$\min\{\tau(B'_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_1)\} \geq \max\{\tau(B'_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_2)\}.$$

Note that $\nu_{d_1}(T) \notin \{\mathrm{r}(B_1), \mathrm{r}(B_2)\}$, otherwise one of B'_1 and B'_2 would be empty. The case where $\nu_{d_1}(T) \notin V(B_1) \cup V(B_2)$ is trivial.

Case 1: $\nu_{d_1}(T) \in V(B_1)$. We can assume that $\nu_{d_1}(T) \in V(B_{1,1})$. For this case we have $\mathrm{rd}(B'_1) \leq \mathrm{rd}(B_1)$ and $\mathrm{rd}(B'_2) = \mathrm{rd}(B_2)$.

If $\mathrm{d}(\nu_{d_1}(T), \mathrm{rd}(B_{1,1}))$ is even, then it follows from $i)$ that

$$\tau(B_{1,1} - \langle \nu_{d_1}(T) \rangle, x) < \tau(B_{1,1}, x).$$

Hence, $a)$ follows from (5.35) and (5.36).

If $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{1,1}))$ is odd (in particular $\nu_{d_1}(T) \neq \mathrm{r}(B_{1,1})$) and there exists $i \in \{1, \ldots, \mathrm{rd}(B_2)\}$ such that $\tau(B_{1,1}, x) = \tau(B_{2,i}, x)$ which also means $B_{1,1} \approx_r B_{2,i}$, then $\mathrm{rd}(B'_1) = \mathrm{rd}(B_1)$, $\mathrm{rd}(B'_2) = \mathrm{rd}(B_2)$ and by $i)$ we have

$$\tau(B_{1,1}, x) < \tau(B_{1,1} - \langle \nu_{d_1}(T) \rangle, x) = \tau(B'_{1,1}, x). \tag{5.37}$$

We can use Lemma 5.17 to obtain $\mathrm{rd}(B'_1) = \mathrm{rd}(B_1) \geq \mathrm{rd}(B_2) = \mathrm{rd}(B'_2)$ and

$$\begin{aligned} \min\{\tau(B'_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_1)\} &\geq \min\{\tau(B_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B_1)\} \\ &\geq \max\{\tau(B_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B_2)\} \\ &= \max\{\tau(B'_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_2)\} \end{aligned}$$

as in $b)$. These are not in contradiction with (5.35) and (5.36), it just means that $\mathrm{rd}(B_1) = \mathrm{rd}(B_2)$ and

$$\tau(B_{1,1}, x) = \cdots = \tau(B_{1,\mathrm{rd}(B_1)}, x) = \tau(B_{2,1}, x) = \cdots = \tau(B_{2,\mathrm{rd}(B_2)}, x).$$

Otherwise, for all $i \in \{1, \dots, \mathrm{rd}(B_2)\}$ we have $\tau(B_{1,1}, x) < \tau(B_{2,i}, x)$. By $ii)$ this implies $\tau(B'_{1,1}, x) \leq \tau(B'_{2,i}, x)$. Therefore $a)$ follows from (5.35) and (5.36).

Case 2: $\nu_{d_1}(T) \in V(B_2)$, say $\nu_{d_1}(T) \in V(B_{2,1})$.

For odd $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{2,1}))$ we have $\tau(B_{2,1} - \langle \nu_{d_1}(T) \rangle, x) > \tau(B_{2,1}, x)$ by $i)$, $\mathrm{rd}(B'_1) = \mathrm{rd}(B_1)$ and $\mathrm{rd}(B'_2) = \mathrm{rd}(B_2)$. Hence, $a)$ follows from (5.35) and (5.36).

If $\mathrm{d}(\nu_{d_1}(T), \mathrm{r}(B_{2,1}))$ is even and there exists $i \in \{1, \dots, \mathrm{rd}(B_1)\}$ such that $\tau(B_{2,1}, x) = \tau(B_{1,i}, x)$, then $\mathrm{rd}(B'_1) = \mathrm{rd}(B_1)$, $\mathrm{rd}(B'_2) \leq \mathrm{rd}(B_2)$ and by $i)$ we have $\tau(B_{2,1}, x) > \tau(B_{2,1} - \langle \nu_{d_1}(T) \rangle, x)$. We can apply Lemma 5.17 to deduce that $\mathrm{rd}(B'_1) = \mathrm{rd}(B_1) \geq \mathrm{rd}(B_2) \geq \mathrm{rd}(B'_2)$ and

$$\begin{aligned} \min\{\tau(B'_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_1)\} &= \min\{\tau(B_{1,i}, x) : 1 \leq i \leq \mathrm{rd}(B_1)\} \\ &\geq \max\{\tau(B_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B_2)\} \\ &\geq \max\{\tau(B'_{2,i}, x) : 1 \leq i \leq \mathrm{rd}(B'_2)\} \end{aligned}$$

as in $b)$.

Otherwise, for all $i \in \{1, \dots, \mathrm{rd}(B_1)\}$ we have $\tau(B_{2,1}, x) > \tau(B_{1,i}, x)$. By $ii)$ this implies $\tau(B'_{2,1}, x) \geq \tau(B'_{1,i}, x)$. Therefore $a)$ follows from (5.35) and (5.36). $\square$

We now have enough tools to prove the first theorem in this chapter. It reveals that there are no non-isomorphic exchange-x-minimal trees with the same degree sequence. This in particular means that exchange-x-minimality and x-minimality are equivalent.

Theorem 5.19 *Let $x > 0$. If T and T' are two exchange-x-minimal trees with the same degree sequence, then T and T' are isomorphic.*

Proof. Let $(d_1, \dots, d_n)$ be the reduced degree sequence of T and T'. We reason by induction with respect to n.

For $n \leq d_n + 1$, it follows from Lemma 5.13 that T and T' have to be isomorphic: each one of them has the form depicted in Figure 5.3 (a). Assume that the theorem is true for $n = k \geq d_n + 1$ and let us show that it holds for $n = k + 1$. Lemma 5.18 shows that $T - \langle \nu_{d_1}(T) \rangle$ and $T' - \langle \nu_{d_1}(T') \rangle$ are exchange-x-minimal trees with the same reduced degree sequence ($(d_2, \dots, d_{k+1} - 1 = d_n - 1)$ or $(d_2, \dots, d_k = d_{n-1})$). By the induction hypothesis there is an isomorphism

$$f_1 : T - \langle \nu_{d_1}(T) \rangle \to T' - \langle \nu_{d_1}(T') \rangle.$$

We will only have to show that $f_1(\nu_{d_n}(T)) = \nu_{d_n}(T')$ for each of the following cases, this implies clearly that f_1 can be extended to an isomorphism $f : T \to T'$, where $f(\nu_{d_1}(T)) = \nu_{d_1}(T')$.

If $d_n \geq 3$, then $\nu_{d_n}(T)$ is the unique vertex of degree $d_n - 1$ in $T - \langle \nu_{d_1}(T) \rangle$; therefore it must be mapped to the unique vertex of degree

$d_n - 1$ in $T' - \langle \nu_{d_1}(T') \rangle$, which is $\nu_{d_n}(T')$. This means that $f_1(\nu_{d_n}(T)) = \nu_{d_n}(T')$.

For the rest of the proof we assume that $d_n = 2$. Hence, in $T - \langle \nu_{d_1}(T) \rangle$ and $T' - \langle \nu_{d_1}(T') \rangle$ respectively, the vertices $\nu_{d_n}(T)$ and $\nu_{d_n}(T')$ are leaves adjacent to $\nu(T)$ and $\nu(T')$. We will only have to show that $f_1(\nu(T)) = \nu(T')$. Because once we have this, even if $f_1(\nu_{d_n}(T)) \neq \nu_{d_n}(T')$, we can define a new isomorphism

$$\begin{aligned} f_1' : T - \langle \nu_{d_1}(T) \rangle &\longrightarrow T' - \langle \nu_{d_1}(T') \rangle \\ v &\longmapsto \begin{cases} f_1(v) & \text{if } v \notin \{f_1^{-1}(\nu_{d_n}(T')), \nu_{d_n}(T)\} \\ f_1(\nu_{d_n}(T)) & \text{if } v = f_1^{-1}(\nu_{d_n}(T')) \\ \nu_{d_n}(T') & \text{if } v = \nu_{d_n}(T) \ , \end{cases} \end{aligned}$$

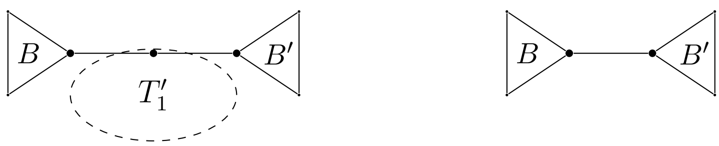
which satisfies $f_1'(\nu_{d_n}(T)) = \nu_{d_n}(T')$ and can therefore be extended to an isomorphism $f : T \to T'$ satisfying $f(\nu_{d_1}(T)) = \nu_{d_1}(T')$.

If $\nu(T)$ is adjacent to at least three non-leaves, then in $T - \langle \nu_{d_1}(T) \rangle$ the vertex $\nu(T)$ becomes the unique non-pseudo-leaf vertex adjacent to a leaf. Necessarily, $f_1(\nu(T))$ is the unique non-pseudo-leaf vertex adjacent to a leaf in $T' - \langle \nu_{d_1}(T') \rangle$. Hence, we must have $f_1(\nu(T)) = \nu(T')$, see Remark 5.16.

For the rest of the proof we assume that each of $\nu(T)$ and $\nu(T')$ has exactly two non-leaf neighbors. Knowing that $d_n = 2$, this means that in $T - \langle \nu_{d_1}(T) \rangle$ and $T' - \langle \nu_{d_1}(T') \rangle$ the vertices $\nu(T)$ and $\nu(T')$, respectively, become pseudo-leaves.

Assume that $f_1(\nu(T)) \neq \nu(T')$. We know that $f_1(\nu(T))$ and $\nu(T')$ are two pseudo-leaves of the same degree in $T' - \langle \nu_{d_1}(T') \rangle$. Let B be the largest complete branch in $T' - \langle \nu_{d_1}(T') \rangle$ containing $f_1(\nu(T))$ such that there exists a complete branch B' in $T' - \langle \nu_{d_1}(T') \rangle$, disjoint from B, which contains $\nu(T')$ and such that there exists an isomorphism $f_2 : B \to B'$ satisfying $f_2(f_1(\nu(T))) = \nu(T')$ and $f_2(\mathrm{r}(B)) = \mathrm{r}(B')$.

If $T' - \langle \nu_{d_1}(T') \rangle$ can be decomposed as in one of the two cases in Figure 5.8, then the function $f_3 : T' - \langle \nu_{d_1}(T') \rangle \to T' - \langle \nu_{d_1}(T') \rangle$ defined



Case 1 Case 2

Figure 5.8: Decomposition of $T' - \langle \nu_{d_1}(T') \rangle$ for the proof of Theorem 5.19

by

$$f_3(v) = \begin{cases} v \text{ if } v \in V(T_1') \\ f_2(v) \text{ if } v \in V(B) \\ f_2^{-1}(v) \text{ if } v \in V(B') \end{cases}$$

is an isomorphism (case 2 corresponds to an empty $V(T_1')$). Since

$$\begin{aligned}(f_3 \circ f_1)(\nu(T)) &= f_3(f_1(\nu(T))) \\ &= f_2(f_1(\nu(T))) \text{ because } f_1(\nu(T)) \in V(B) \\ &= \nu(T'),\end{aligned}$$

$f_3 \circ f_1 : T - \langle \nu_{d_1}(T) \rangle \to T' - \langle \nu_{d_1}(T') \rangle$ can be extended to be an isomorphism from T to T'.

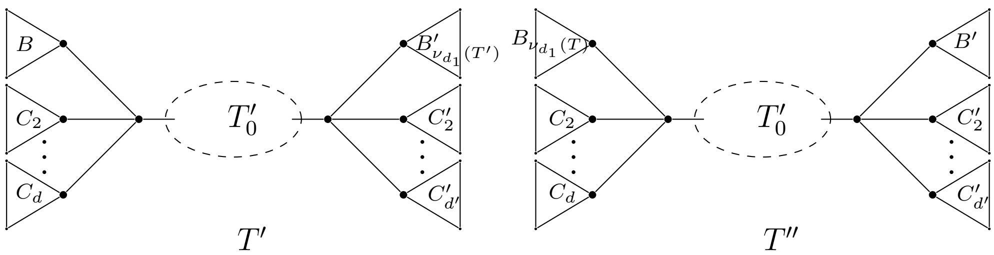
Otherwise, there exist two disjoint complete branches $C = [C_1 = B, C_2, \ldots, C_{\mathrm{rd}(C)}]$ and $C' = [C_1' = B', C_2', \ldots, C'_{\mathrm{rd}(C')}]$ in $T' - \langle \nu_{d_1}(T') \rangle$. By swapping notations if necessary, we can assume that $\mathrm{rd}(C) \leq \mathrm{rd}(C')$ and

$$\max\{\tau(C_i, x) : 1 \leq i \leq \mathrm{rd}(C)\} \leq \min\{\tau(C_i', x) : 1 \leq i \leq \mathrm{rd}(C')\}.$$

It is impossible to have $\mathrm{rd}(C) = \mathrm{rd}(C')$ and $C_2 \approx_r \cdots \approx_r C_{\mathrm{rd}(C)} \approx_r C_2' \approx_r \cdots \approx_r C'_{\mathrm{rd}(C')}$, because this would allow us to extend f_2 to be an isomorphism between C and C', which contradicts the choice of B to be largest. Hence there are $i_0 \geq 2$ and $j_0 \geq 2$ such that

$$\tau(C_{i_0}, x) < \tau(C'_{j_0}, x), \text{ or } \mathrm{rd}(C) < \mathrm{rd}(C'). \tag{5.38}$$

Let $B'_{\nu_{d_1}(T')}$ be the complete branch of T' obtained by returning $\nu_{d_1}(T')$ and its leaves to B'. Note that a graph isomorphic to T can be obtained from $T' - \langle \nu_{d_1}(T') \rangle$ by joining a leaf neighbor of $f_1(\nu(T))$ to the center of a star S_{d_1} by an edge. Let $B_{\nu_{d_1}(T)}$ be the complete branch obtained from B by joining a leaf of $f_1(\nu(T))$ to the center of a star S_{d_1}. Clearly we have $B_{\nu_{d_1}(T)} \approx_r B'_{\nu_{d_1}(T')}$, by extension of f_2. Now we have the decompositions:



where T'' is isomorphic to T. If $\tau(B, x) < \tau(B_{\nu_{d_1}(T)}, x) = \tau(B'_{\nu_{d_1}(T')}, x) > \tau(B', x)$, then combined with (5.38) this contradicts the exchange-x-minimality of T. Otherwise we have

$$\tau(B, x) > \tau(B_{\nu_{d_1}(T)}, x) = \tau(B'_{\nu_{d_1}(T')}, x) < \tau(B', x),$$

in view of (5.38) this contradicts the exchange-x-minimality of T'. □

Corollary 5.20 *For all $x > 0$, a tree T is exchange-x-minimal if and only if it is x-minimal.*

Corollary 5.21 *For all $x > 0$, if T and T' are two x-minimal trees with the same degree sequence, then T is isomorphic to T'.*

5.3 Characterization of the x-minimal tree

We will see in this section how to construct an x-minimal tree corresponding to any given degree sequence. But before this, one more lemma is still needed. It describes a way to remove some leaves without losing the x-minimality of a tree.

Lemma 5.22 *Let $x > 0$, let T be a tree with reduced degree sequence $(d_1, \ldots, d_n)$, where $n \geq d_n + 2$, and with $d_n + k$ pseudo-leaves for some integer $k \geq 0$. Assume that for some vertex ν, T can be decomposed as on the left hand side of Figure 5.9, where the subgraph T_0 does not contain any leaves or pseudo-leaves of T. Let T^+ be the tree obtained from T by attaching a leaf to each pseudo-leaf of T and to ν, see the right hand side of Figure 5.9. If T^+ is an x-minimal tree, then so is T.*

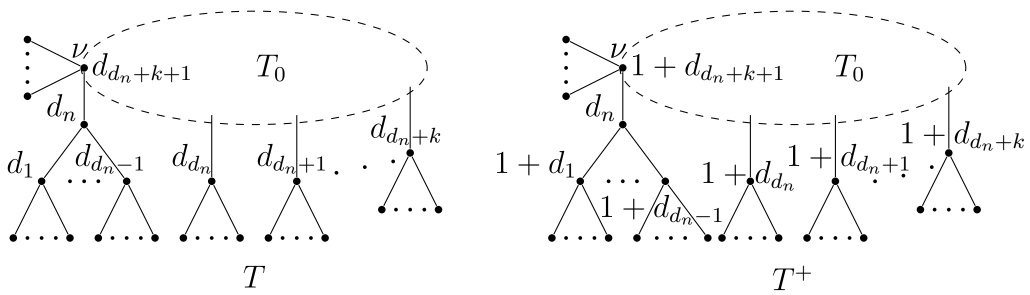


Figure 5.9: Construction of T^+

Proof. The reduced degree sequence of T^+ is

$$(1 + d_1, \ldots, 1 + d_{d_n+k+1}, d_{d_n+k+2}, \ldots, d_n).$$

For all complete branches B of T^+, B^- denotes the corresponding complete branch of T obtained from B by removing the additional leaves which are not vertices of T. We assume that T^+ is x-minimal.

Note that in T, as well as in T^+, leaves are all attached to vertices with largest degrees. Moreover, since ν is a non-pseudo-leaf vertex adjacent to a leaf in T^+, by Lemma 5.12 and Remark 5.16 we have $\nu = \nu(T^+)$.

To see that the exchange-x-minimality of T^+ implies exchange-x-minimality of T, we are left to show that for any two disjoint non-leaf complete branches B and B' in T^+ we have the following:

i) $\tau(B, x) = \tau(B', x)$ implies $\tau(B^-, x) = \tau(B'^-, x)$,

ii) $\tau(B, x) < \tau(B', x)$ implies $\tau(B^-, x) \leq \tau(B'^-, x)$.

The exchange-x-minimality of T^+, *i)* and *ii)* imply exchange-x-minimality of T. Then the lemma follows from Corollary 5.20.

Assume that

$$\tau(B, x) = \tau(B', x), \tag{5.39}$$

then by Lemma 5.14 it follows that

$$B \approx_r B'. \tag{5.40}$$

An isomorphism maps non-pseudo-leaf vertices to non-pseudo-leaf vertices of the same degree. Since $\nu(T^+)$ is the unique non-pseudo-leaf vertex of degree $1 + d_{d_n+k+1}$ in T^+, we deduce that $\nu(T^+)$ is not in $V(B) \cup V(B')$. This means that in B and B' the additional leaves are adjacent to pseudo-leaves. Therefore (5.40) implies that $B^- \approx_r B'^-$, and consequently $\tau(B^-, x) = \tau(B'^-, x)$.

To see *ii)*, we reason by induction with respect to

$$M_h := \max\{\mathrm{h}(B), \mathrm{h}(B')\}.$$

If $M_h = 1$, then $\mathrm{r}(B)$ and $\mathrm{r}(B')$ are pseudo-leaves. Thus, $\tau(B, x) < \tau(B', x)$ implies that $\mathrm{rd}(B^-) = \mathrm{rd}(B) - 1 > \mathrm{rd}(B') - 1 = \mathrm{rd}(B'^-)$ and hence $\tau(B^-, x) < \tau(B'^-, x)$.

Assume that *ii)* holds if $M_h = j$ for some positive integer j. Now let B and B' be such that $M_h = \max\{h(B), h(B')\} = j + 1$ and

$$\tau(B, x) < \tau(B', x). \tag{5.41}$$

By Lemma 5.10, we also have

$$\mathrm{rd}(B) \geq \mathrm{rd}(B'). \tag{5.42}$$

If $\mathrm{h}(B) = 1$ and $\mathrm{h}(B') \geq 2$, then it follows that $\mathrm{rd}(B^-) \geq \mathrm{rd}(B'^-)$ (remember that in T and in T^+ the degree of a non-pseudo-leaf vertex is less or equal to the degree of any pseudo-leaf, see Figure 5.9.) and

$$\tau(B^-, x) = \tau([1 + \mathrm{rd}(B^-)], x) \leq \tau([1 + \mathrm{rd}(B'^-)], x) < \tau(B'^-, x).$$

It is impossible to have $\mathrm{h}(B) \geq 2$ and $\mathrm{h}(B') = 1$, since it leads to $\mathrm{rd}(B) \leq \mathrm{rd}(B')$ and $\tau(B', x) = \tau([1 + \mathrm{rd}(B')], x) \leq \tau([1 + \mathrm{rd}(B)], x) <$

$\tau(B, x)$, and thus it contradicts (5.41). For the rest of the proof we can assume that $\min\{\mathrm{h}(B_1), \mathrm{h}(B_2)\} \geq 2$. Let $B = [B_1, \dots, B_{\mathrm{rd}(B)}]$ and $B' = [B'_1, \dots, B'_{\mathrm{rd}(B')}]$. Using Lemma 5.10, (5.41) implies

$$\min\{\tau(B_i, x) : 1 \leq i \leq \mathrm{rd}(B)\} \geq \max\{\tau(B'_i, x) : 1 \leq i \leq \mathrm{rd}(B')\}. \quad (5.43)$$

With the induction hypothesis and $i)$ we also get

$$\min\{\tau(B_i^-, x) : 1 \leq i \leq \mathrm{rd}(B^-)\} \geq \max\{\tau(B_i'^-, x) : 1 \leq i \leq \mathrm{rd}(B'^-)\}. \quad (5.44)$$

If $\nu \neq \mathrm{r}(B), \mathrm{r}(B')$, then we have $\mathrm{rd}(B^-) = \mathrm{rd}(B) \geq \mathrm{rd}(B') = \mathrm{rd}(B'^-)$. With (5.44) this leads to

$$\tau(B^-, x) = \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B^-)} \tau(B_i^-)} \leq \frac{1}{1 + x\sum_{i=1}^{\mathrm{rd}(B'^-)} \tau(B_i'^-, x)} = \tau(B'^-, x).$$

Since the degree of ν, as a vertex of T^+, is strictly greater than the degree of any non-pseudo-leaf vertex of T^+, it is only $\mathrm{r}(B)$ which can possibly be ν, but not $\mathrm{r}(B')$. Assume that $\nu = \mathrm{r}(B)$. Then $\mathrm{rd}(B) > \mathrm{rd}(B')$ and thus

$$\mathrm{rd}(B^-) = \mathrm{rd}(B) - 1 \geq \mathrm{rd}(B') = \mathrm{rd}(B'^-).$$

Combined with (5.44) this gives $\tau(B^-, x) \leq \tau(B'^-, x)$ again. $\square$

Now we describe a tree which will be shown later to be x-minimal.

Definition 5.23 Let $(d_1, \dots, d_n)$ be a reduced degree sequence of a tree. If $n \leq d_n + 1$, then $\mathcal{M}(d_1, \dots, d_n)$ is the tree obtained by merging the root of each of $[d_1], \dots, [d_{n-1}]$ with a leaf of $[1+d_n]$, respectively. We label selected vertices as shown in Figure 5.10, in such a way that

$$\deg(v_i) \leq \deg(v_j) \text{ if } i < j. \quad (5.45)$$

At this stage all non-leaf vertices are labeled. If $n \geq d_n + 2$, we con-

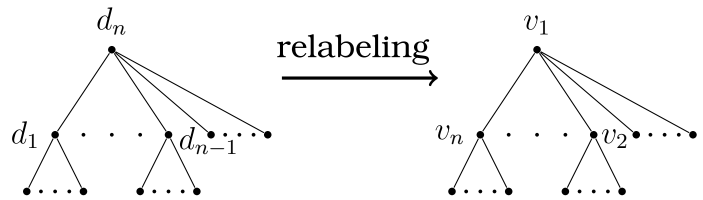


Figure 5.10: Relabeling for the construction of $\mathcal{M}(d_1, \dots, d_n)$

struct $\mathcal{M}(d_1, \dots, d_n)$ recursively: let l be the largest integer such that v_l is a label in $\mathcal{M}(d_{d_n}, \dots, d_{n-1})$. Let s be the smallest integer such that v_s is adjacent to a leaf in $\mathcal{M}(d_{d_n}, \dots, d_{n-1})$. Let $R_{d_n} = [[d_1], \dots, [d_{d_n-1}]]$, where the pseudo-leaves are labeled $v_{l+1}, \dots, v_{l+d_n-1}$ still respecting (5.45). $\mathcal{M}(d_1, \dots, d_n)$ is the tree obtained by merging the root of R_{d_n} with a leaf adjacent to v_s. See Figure 5.11 for an example.

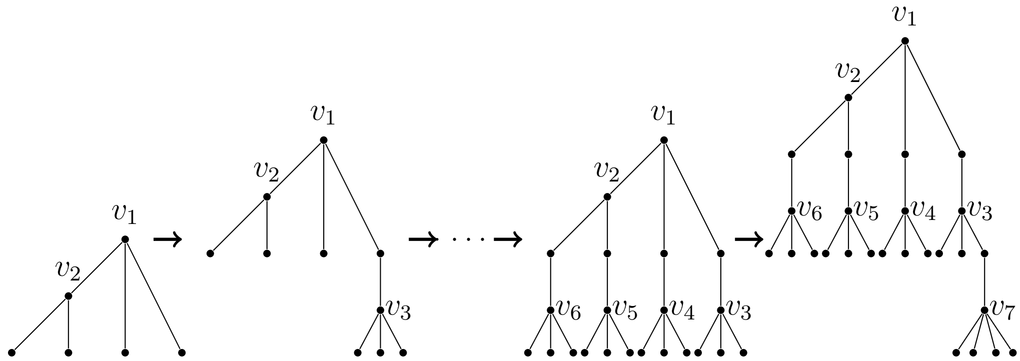


Figure 5.11: Step-by-step construction of $\mathcal{M}(5, 4, 4, 4, 4, 3, 3, 2, 2, 2, 2, 2)$

The main theorem of this chapter reads as follows:

Theorem 5.24 *Let $x > 0$. If T is a tree with reduced degree sequence $(d_1, \ldots, d_n)$ that is not isomorphic to $\mathcal{M}(d_1, \ldots, d_n)$, then we have*

$$\mathrm{M}(T, x) > \mathrm{M}(\mathcal{M}(d_1, \ldots, d_n), x).$$

Proof. Note that the tree $\mathcal{M}(d_1, \ldots, d_n)$ has all the properties stated in Lemma 5.12. The proof is by induction with respect to n. For $n \leq d_n + 1$, we have already seen in Lemma 5.13 that the structure of the x-minimal tree coincides with that of $\mathcal{M}(d_1, \ldots, d_n)$. Assume that the theorem holds whenever $n \leq k$, for some integer k at least equal to d_n. Now, assume that $n = k + 1$. Let j be the number of pseudo-leaves in the x-minimal tree with reduced degree sequence $(d_1, \ldots, d_n)$. First we show that

$$T_M = \mathcal{M}(1 + d_1, \ldots, 1 + d_{j+1}, d_{j+2}, \ldots, d_n)$$

is x-minimal. We know (by Lemma 5.12 and Remark 5.16) that the x-minimal tree, say H, with reduced degree sequence $(1 + d_1, \ldots, 1 + d_{j+1}, d_{j+2}, \ldots, d_n)$ can be represented as in Figure 5.12, for some tree

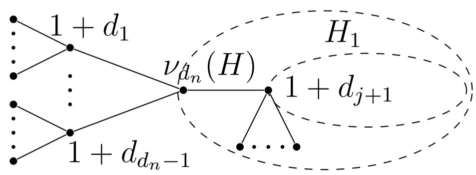


Figure 5.12: Decomposition of the tree H in the proof of Theorem 5.24

H_1. Applying (3.17) and (3.19) repeatedly to edges incident to $\nu_{d_n}(H)$

and not contained in H_1 we obtain

$$\mathrm{M}(H,x) = \mathrm{M}(H_1,x)\prod_{i=1}^{d_n-1}\mathrm{M}([1+d_i],x) + x\,\mathrm{M}(H_1-\nu_{d_n}(H),x)\sum_{l=1}^{d_n-1}\prod_{\substack{i=1\\ i\neq l}}^{d_n-1}\mathrm{M}([1+d_i],x). \tag{5.46}$$

On the other hand, as seen in Definition 5.23, T_M also has the form shown in Figure 5.12 if we replace H_1 by $\mathcal{M}(1+d_{d_n},\ldots,1+d_{j+1},d_{j+2},\ldots,d_{n-1})$ and thus $H_1-\nu_{d_n}(H)$ corresponds to $\mathcal{M}(1+d_{d_n},\ldots,1+d_j,d_{j+1},\ldots,d_{n-1})$, since reducing the degree $d_{j+1}+1$ by 1 does not change the order of the degrees. The representation in Figure 5.12 also shows that T_M and H both have j pseudo-leaves. We get

$$\begin{aligned}
&\mathrm{M}(T_M,x)\\
&= \mathrm{M}(\mathcal{M}(1+d_{d_n},\ldots,1+d_{j+1},d_{j+2},\ldots,d_{n-1}),x)\prod_{i=1}^{d_n-1}\mathrm{M}([1+d_i],x)\\
&+ x\,\mathrm{M}(\mathcal{M}(1+d_{d_n},\ldots,1+d_j,d_{j+1},\ldots,d_{n-1}),x)\sum_{l=1}^{d_n-1}\prod_{\substack{i=1\\ i\neq l}}^{d_n-1}\mathrm{M}([1+d_i],x).
\end{aligned} \tag{5.47}$$

Using the induction hypothesis we can deduce by comparing (5.46) and (5.47) that $\mathrm{M}(T_M,x) \le \mathrm{M}(H,x)$. Since H is x-minimal, equality has to hold. By the uniqueness of the x-minimal tree (see Theorem 5.19), H and T_M must be isomorphic. Finally, the x-minimality of $\mathcal{M}(d_1,d_2,\ldots,d_n)$ follows from Lemma 5.22. $\square$

Corollary 5.25 *If a tree T, with reduced degree sequence $(d_1,\ldots,d_n)$, is not isomorphic to $\mathcal{M}(d_1,\ldots,d_n)$, then we have*

$$\mathrm{En}(T) > \mathrm{En}(\mathcal{M}(d_1,\ldots,d_n))$$

and $\mathrm{Z}(T) > \mathrm{Z}(\mathcal{M}(d_1,\ldots,d_n))$.

Proof. See relations (3.15) and (3.16). $\square$

For any rooted tree R let $\rho(R) = \sigma_0(R)/\sigma(R)$, where $\sigma_0(R)$ is the number of independent vertex subsets in R not covering the root. An exchange lemma with respect to ρ is obtained in [42]. A degree sequence preserving version of the lemma can also be deduced, similar to Lemma 5.9. From then, the same process as above with minor adjustments can be used to prove the uniqueness (up to isomorphism) of the tree with given degree sequence and maximum

Merrifield-Simmons index. Furthermore, it turns out that the graph transformations described in Lemmas 5.18 and 5.22 also preserve the maximality of Merrifield-Simmons index among all trees of the same degree sequence. Hence we obtain the following theorem:

Theorem 5.26 *$\mathcal{M}(d_1, \ldots, d_n)$ is the unique (up to isomorphism) tree with reduced degree sequence $(d_1, \ldots, d_n)$ and maximum Merrifield-Simmons index.*

5.4 Comparing x-minimal trees with different degree sequences

For any degree sequence $D = (d_1, \ldots, d_n, 1, \ldots, 1)$, where $d_n \geq 2$, we define $\mathcal{M}(D) = \mathcal{M}(d_1, \ldots, d_n)$.

If $(d_1, \ldots, d_n)$ *majorizes* $(b_1, \ldots, b_n)$ and there exists $i_0 \in \{1, \ldots, n\}$ such that $d_{i_0} \neq b_{i_0}$, then we write

$$(b_1, \ldots, b_n) \prec (d_1, \ldots, d_n).$$

The following lemma follows directly from Lemma 5.8.

Lemma 5.27 (cf. Lemma 3.1 of [43]) *Let $x > 0$, and let T be an x-minimal tree. Let $B = [B_1, \ldots, B_{\mathrm{rd}(B)}]$ and $B' = [B'_1, \ldots, B'_{\mathrm{rd}(B')}]$ be two disjoint complete non-leaf branches of T such that $\mathrm{rd}(B) \leq \mathrm{rd}(B')$ and*

$$\max\{\tau(B_i, x) : 1 \leq i \leq \mathrm{rd}(B)\} \leq \min\{\tau(B'_i, x) : 1 \leq i \leq \mathrm{rd}(B')\}.$$

For the case where $B \approx_r B'$, we swap notations if needed so that

$$\mathrm{M}(T_0 - \{\mathrm{r}(B), \mathrm{r}(B'), \mathrm{rn}(B')\}) \geq \mathrm{M}(T_0 - \{\mathrm{rn}(B), \mathrm{r}(B), \mathrm{r}(B')\}),$$

an inequality that has to be satisfied by application of Lemma 5.9 if $B \not\approx_r B'$. Let T' be the tree obtained from T by replacing B and B' by $C = [C_1 = B_2, \ldots, C_{\mathrm{rd}(C)} = B_{\mathrm{rd}(B)}]$ and $C' = [C'_1 = B_1, C'_2 = B'_1, \ldots, C'_{\mathrm{rd}(C')} = B'_{\mathrm{rd}(B')}]$, respectively (see Figure 5.13). Then we have $\mathrm{M}(T', x) < \mathrm{M}(T, x)$.

Proof. Using the same notation as in the proof of Lemma 5.8 but

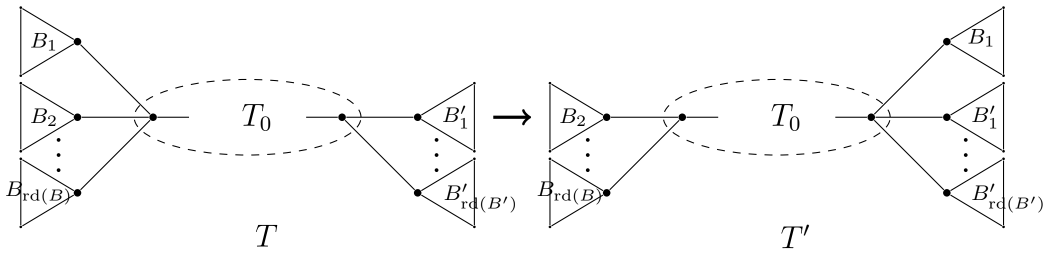


Figure 5.13: Construction of the tree T' in Lemma 5.27

replacing w, w', v', v by $\mathrm{r}(B), \mathrm{rn}(B), \mathrm{rn}(B'), \mathrm{r}(B')$ respectively, we have

$$\begin{aligned}
&\frac{\mathrm{M}(T,x)-\mathrm{M}(T',x)}{\prod_{i=1}^{\mathrm{r}(B)}\mathrm{M}(B_i,x)\prod_{i=1}^{\mathrm{r}(B')}\mathrm{M}(B'_i,x)}\\
&= (G_{T_0}(B_2,\dots,B_{\mathrm{rd}(B)};B_1,B'_1,\dots,B'_{\mathrm{rd}(B')})\\
&\qquad\qquad\qquad - G_{T_0}(B_1,\dots,B_{\mathrm{rd}(B)};B'_1,\dots,B'_{\mathrm{rd}(B')}))\\
&= \left(\sum_{i=1}^{\mathrm{rd}(B')}\tau(B'_i,x)-\sum_{i=2}^{\mathrm{rd}(B)}\tau(B_i,x)\right)x^2\tau(B_1,x)\,\mathrm{M}\left(T_0-\{\mathrm{r}(B),\mathrm{r}(B')\},x\right)\\
&+\Big(\mathrm{M}(T_0-\{\mathrm{r}(B),\mathrm{r}(B'),\mathrm{rn}(B')\})-\mathrm{M}(T_0-\{\mathrm{rn}(B),\mathrm{r}(B),\mathrm{r}(B')\})\Big)x^2\tau(B_1,x)\\
&> 0.
\end{aligned}$$

□

Theorem 5.28 *Let $(b_1,\dots,b_n)$ and $(d_1,\dots,d_n)$ be two degree sequences of trees of the same length. If $(b_1,\dots,b_n)\prec(d_1,\dots,d_n)$, then for all $x>0$ we have*

$$\mathrm{M}(\mathcal{M}(b_1,\dots,b_n),x) > \mathrm{M}(\mathcal{M}(d_1,\dots,d_n),x).$$

Proof. Assume that $B=(b_1,\dots,b_n)$, $D=(d_1,\dots,d_n)$ are degree sequences of trees and $B\prec D$. Then,

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} d_i, \tag{5.48}$$

and there exists i_0 such that $d_{i_0}\neq b_{i_0}$. In fact the set $\mathbb{I}=\{i: d_i\neq b_i\}$ must have at least two elements, otherwise (5.48) would be impossible. Let $l=\min\{i: d_i\neq b_i\}$ and $m=\max\{i: d_i\neq b_i\}$. We must have $b_l<d_l$ and $b_m>d_m$. We define

$$B_1=(b_1,\dots,b_{l-1},b_l+1,b_{l+1},\dots,b_{m-1},b_m-1,b_{m+1},\dots,b_n).$$

Note that $b_{l-1}=d_{l-1}\geq d_l\geq b_l+1$ and $b_{m+1}=d_{m+1}\leq d_m\leq b_m-1$, which means that B_1 is a valid degree sequence. It is easy to see that $B \prec B_1$. Most importantly, by application of Lemma 5.27 to $\mathcal{M}(B)$, we know that there exists a tree T with degree sequence B_1 obtained by moving a branch in $\mathcal{M}(B)$ from a vertex of degree b_m to a vertex of degree b_l, and such that

$$\mathrm{M}(\mathcal{M}(B), x) > \mathrm{M}(T, x) \geq \mathrm{M}(\mathcal{M}(B_1), x).$$

If $B_1 = D$, then we are done. Otherwise, we iterate the process. We set $B = B_0$, and if k is a positive integer and $B_k \neq D$, then we construct B_{k+1} using exactly the same way as B_1 was constructed from B. After a finite number $K = \frac{1}{2}\sum_{i\in\mathbb{I}} |d_i - b_i|$ of iterations we will get the chain

$$B = B_0 \prec B_1 \prec \cdots \prec B_{K-1} \prec B_K = D.$$

For any $k \in \{1, \ldots, K-1\}$, we can apply Lemma 5.27 to $\mathcal{M}(B_k)$ to deduce that there exists a tree T_{k+1} with degree sequence B_{k+1} and such that

$$\mathrm{M}(\mathcal{M}(B_k), x) > \mathrm{M}(T_{k+1}, x) \geq \mathrm{M}(\mathcal{M}(D_{k+1}), x).$$

In total we get

$$\mathrm{M}(\mathcal{M}(B), x) > \mathrm{M}(\mathcal{M}(B_1), x) > \cdots > \mathrm{M}(\mathcal{M}(B_K), x) = \mathrm{M}(\mathcal{M}(D), x).$$

This completes the proof. □

Using a version of Lemma 5.27 with respect to the Merrifield-Simmons index, one also obtains the following:

Theorem 5.29 *Let $(b_1, \ldots, b_n)$ and $(d_1, \ldots, d_n)$ be two degree sequences. If $(b_1, \ldots, b_n) \prec (d_1, \ldots, d_n)$, then $\sigma(\mathcal{M}(d_1, \ldots, d_n)) > \sigma(\mathcal{M}(b_1, \ldots, b_n))$.*

For a class of graphs C, we denote by $\mathrm{Extr}(C)$ the element of C which has minimum energy and Hosoya index and maximum Merrifield-Simmons index, if such a tree exists and if it is unique (up to isomorphism). The following known results now follow as simple corollaries of Theorems 5.28 and 5.29. In each corollary, we denote by $D(\mathrm{Extr}(C))$ and $D(T)$ the degree sequences of $\mathrm{Extr}(C)$ and a given $T \in C$, respectively, and one can always see that $D(T) \preccurlyeq D(\mathrm{Extr}(C))$.

Corollary 5.30 ([41, 74]) *Let C be the class of n-vertex trees, then*

$$\mathrm{Extr}(C) = \mathcal{M}(n-1, 1, \ldots, 1).$$

Corollary 5.31 ([42, 43]) *Let C be the class of n-vertex trees whose maximum degree is d. Write $n - 2 = k(d-1) + r$, for non-negative integers k and $r < d - 1$. Then we have $\mathrm{Extr}(C) = \mathcal{M}(d_1, \ldots, d_n)$, where $d_1 = \cdots = d_k = d$, $d_{k+1} = r + 1$ and $d_{k+2} = \cdots = d_n = 1$. This means that*

$$\mathrm{Extr}(C) = \mathcal{M}(d, \ldots, d, r+1, 1 \ldots, 1).$$

Proof. The tree $\mathrm{Extr}(C)$ described in the corollary has as many vertices of degree d as possible for an element of C. It is now easy to see that if T is an other element of C and $D(T) \neq D(\mathrm{Extr}(C))$, then we have $D(T) \prec D(\mathrm{Extr}(C))$. □

Corollary 5.32 ([73, 88]) *Let C be the class of n-vertex trees with exactly d leaves. Then we have $\mathrm{Extr}(C) = \mathcal{M}(d_1, \ldots, d_n)$, where $d_1 = d$, $d_2 = \cdots = d_{n-d} = 2$ and $d_{n-d+1} = \cdots = d_n = 1$. That is,*

$$\mathrm{Extr}(C) = \mathcal{M}(d, 2, \ldots, 2, 1, \ldots, 1).$$

Proof. If $D(T) = (b_1, \ldots, b_n)$ is a degree sequence of an element T of C, then we must have $b_1 \leq d$ and $b_{n-d+1} = \cdots = b_n = 1$, each of the other degrees is at least 2. It is easy to see that for any $n \geq l \geq 1$ we have $\sum_{i=l}^{n} b_i \geq \sum_{i=l}^{n} d_i$. Therefore, if $D(T) \neq D(\mathrm{Extr}(C))$ then we have $D(T) \prec D(\mathrm{Extr}(C))$. □

Corollary 5.33 ([62, 86]) *Let C be the class of n-vertex trees with diameter d. Then we have $\mathrm{Extr}(C) = \mathcal{M}(d_1, \ldots, d_n)$ where*

$$(d_1, \ldots, d_n) = (n - d + 1, 2, \ldots, 2, 1, \ldots, 1),$$

more precisely $d_1 = n - d + 1$, $d_2 = \cdots = d_{d-1} = 2$ and $d_d = \cdots = d_n = 1$.

Proof. Let $D = (d_1, \ldots, d_n)$ be a degree sequence of a tree with diameter $d \leq n - 1$. The degrees of the $d - 1$ non-leaves on a diameter of T are at least 2. It follows that $d_i \geq 2$ for all $i \leq d-1$ (and of course $d_i \geq 1$ for all i), so that D is indeed majorized by $(n - d + 1, 2, \ldots, 2, 1, \ldots, 1)$. □

Chapter 6

Unicyclic graphs with large energy

One of the classes of graphs that has been quite thoroughly studied in the literature on the energy of graphs is the class of unicyclic graphs, i.e., connected graphs with only one unique cycle. It is the most natural class of graphs to consider after that of trees, many of the techniques that apply to trees can still be used for unicyclic graphs. The motivation of studying these graphs is strengthened by the fact that chemical compounds which have unicyclic molecular graphs are not uncommon.

Among all unicyclic graphs with a given number $n \geq 6$ of vertices, the minimum energy is attained for the graph that results from connecting two leaves of a star by an edge; the second-smallest, third-smallest, ..., sixth-smallest values and the corresponding graphs are also all known [12, 50, 65, 69]. On the other hand, the converse question for the largest possible energy of a unicyclic graph appears to be somewhat more intricate. In answering this question, the so-called tadpole graphs P_n^k, which are obtained by merging an end of a path of length $n-k$ with a vertex in a cycle of length k, play an essential role. We define $\mathbb{U}_n$ to be the set of all n-vertex unicyclic graphs, and $\mathbb{U}_n(l)$ the subset of $\mathbb{U}_n$ whose elements are the n-vertex unicyclic graphs whose cycle has length l. The following was originally found by means of an extensive computer search:

Conjecture 6.1 ([11, 38]) *Among all elements of $\mathbb{U}_n$ for any given $n \geq 7$, the cycle C_n has maximal energy if $n = 9, 10, 11, 13$ and 15. For all other values of n the unicyclic graph with maximum energy is P_n^6 (see Figure 6.1).*

Substantial progress on this conjecture was already made shortly afterwards in a paper of Hou, Gutman and Woo [51], who proved the following:

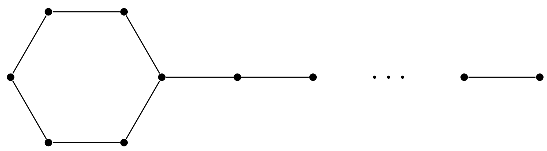


Figure 6.1: The graph P_n^6

Theorem 6.2 ([51]) *Let $C(n,k)$ be the set of all unicyclic graphs obtained from a cycle C_k of length k by adding $n-k$ pendant vertices to it. Suppose that $G \in \mathbb{U}_n(k)$, $n \geq k$. If G has maximum energy in $\mathbb{U}_n(k)$, then G is either P_n^k or, when $k \equiv 0 \mod 4$, a graph from $C(n,k)$.*

By virtue of this theorem, the authors of [51] were also able to prove the following for the slightly narrower class of unicyclic bipartite graphs:

Theorem 6.3 ([51]) *P_n^6 has the largest energy among all unicyclic bipartite n-vertex graphs, except possibly the cycle C_n.*

In [2] we proved that P_n^6 "wins" over C_n for almost all n. The same fact was rediscovered [56] independently by Huo, Li and Shi. The corresponding theorem reads as follows:

Theorem 6.4 ([2, 56]) *For all $n \geq 7$, $n \neq 9, 10, 11, 13, 15$, we have $\mathrm{En}(C_n) < \mathrm{En}(P_n^6)$.*

This completes Theorem 6.3 and proves that P_n^6 is the bipartite unicyclic graph with n vertices and maximum energy, if $n \geq 7$ and $n \neq 9, 10, 11, 13, 15$. The current chapter is a natural continuation of this work. Our proof [6], detailed in the second section, for the more general Conjecture 6.1 appeared almost simultaneously with an other proof of the same conjecture by Huo, Li and Shi in [55]. While the proof given in [55] is simpler in some respects, the main benefit of our method is that it gives more precise information on the actual value of the maximum energy of an n-vertex unicyclic graph and the gap between P_n^6 and other graphs of the form P_n^k (it turns out that, for fixed k, the difference of the energies converges to a constant as $n \to \infty$). We are also able to prove additional results by means of our method:

Theorem 6.5 *Let D_n be the graph obtained by joining a vertex of the cycle C_6 and the third vertex in P_{n-6} by an edge (see Figure 6.2). For $n \geq 28$, D_n is the unicyclic graph with second-largest energy.*

This was conjectured in [34] for bipartite unicyclic graphs, but it still holds in $\mathbb{U}_n$. In order to prove this theorem, we can make use of the following result:

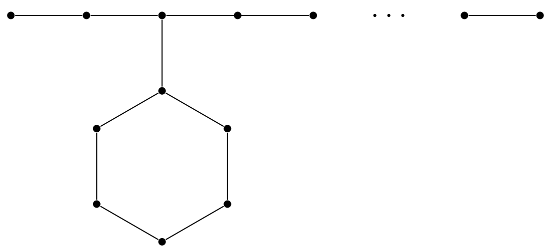


Figure 6.2: The graph D_n

Theorem 6.6 ([52]) *Let* $G \in \mathbb{U}_n \setminus \{P_n^k | k = 3, 4, \dots, n\}$ *be a bipartite unicyclic graph. For* $n \geq 13$, *if* $G \neq D_n$, *we have* $\mathrm{En}(G) < \mathrm{En}(D_n)$.

This shows that it is sufficient to compare D_n to all tadpole graphs P_n^k once Conjecture 6.1 has been verified. In the following section, we gather some auxiliary tools. We then proceed to determine an integral representation for $\mathrm{En}(P_n^k)$, which is used to study the behavior of P_n^k as k varies. This leaves us with only a few cases that are studied in more detail to prove Conjecture 6.1 as well as Theorem 6.5. Another result that we obtain as a consequence of our estimates is the following:

Theorem 6.7 *Among all non-bipartite unicyclic graphs with at least three vertices,* P_n^3 *has maximum energy if* n *is even, and* C_n *has maximum energy if* n *is odd.*

6.1 Preliminaries

As most other results on the energy of graphs, our proofs are based on (3.13). If we write the characteristic polynomial of G as $\Phi(G, x) = \sum_{k=1}^{n} a_k x^{n-k}$, then we have

$$
\begin{aligned}
|\Phi(G, \mathbf{i}/x)|^2 &= \left| \sum_{k=0}^{n} a_k (\mathbf{i}/x)^{n-k} \right|^2 = |\mathbf{i}^n|^2 x^{-2n} \left| \sum_{k=0}^{n} a_k (\mathbf{i}/x)^{-k} \right|^2 \\
&= x^{-2n} \left[\left(\sum_{k \geq 0} (-1)^k a_{2k} x^{2k} \right)^2 + \left(\sum_{k \geq 0} (-1)^k a_{2k+1} x^{2k+1} \right)^2 \right],
\end{aligned}
$$

and hence the integral (3.13) can also be written as

$$
\mathrm{En}(G) = \frac{2}{\pi} \int_0^{+\infty} \frac{1}{x^2} \log \left(x^n \mid \Phi(G, \mathbf{i}/x) \mid \right) dx.
$$

Since
$$\frac{d}{du}\frac{1}{\sinh u} = -\frac{\cosh u}{\sinh^2 u},$$
the change of variable $x = 1/(2\sinh u)$ leads us to

$$\begin{aligned}
\mathrm{En}(G) &= \frac{2}{\pi}\int_0^{+\infty} du \frac{\cosh u}{2\sinh^2 u}(4\sinh^2 u)\log\left(\frac{1}{2^n \sinh^n u}|\Phi(G, 2\mathbf{i}\sinh u)|\right) \\
&= \frac{4}{\pi}\int_0^{+\infty} \log\left(\frac{|\Phi(G, 2\mathbf{i}\sinh u)|}{2^n\sinh^n u}\right)\cosh u\, du, \qquad (6.1)
\end{aligned}$$

which is the expression for the energy that we will mostly work with.

6.2 The energy of the tadpole graph P_n^k

6.2.1 A formula for $\mathrm{En}(P_n^k)$

In order to determine $\mathrm{En}(P_n^k)$ using (6.1), we need an explicit expression for $\Phi(P_n^k, 2\mathbf{i}\sinh u)$. Using Lemma 3.8 we obtain

$$\begin{aligned}
\Phi(P_k, x) &= x\Phi(P_{k-1}, x) - \Phi(P_{k-2}, x), \qquad (6.2)\\
\Phi(C_k, x) &= \Phi(P_k, x) - \Phi(P_{k-2}, x) - 2,\\
\Phi(P_n^{n-1}, x) &= x\Phi(C_{n-1}, x) - \Phi(P_{n-2}, x)
\end{aligned}$$

and for all $k \le n-2$ we have

$$\Phi(P_n^k, x) = x\Phi(P_{n-1}^k, x) - \Phi(P_{n-2}^k, x). \qquad (6.3)$$

Note that $\Phi(P_n^n, x) = \Phi(C_n, x)$. The characteristic equation

$$q^2 - xq + 1 = q^2 - 2\mathbf{i}q\sinh u + 1 = 0 \qquad (6.4)$$

of the linear recurrence (6.2) (taking $x = 2\mathbf{i}\sinh u$) has the two roots

$$q_1 = \frac{x + \sqrt{x^2-4}}{2} = \mathbf{i}e^u \qquad (6.5)$$

and

$$q_2 = \frac{x - \sqrt{x^2-4}}{2} = (\mathbf{i}e^u)^{-1}. \qquad (6.6)$$

Thus, we have an explicit formula of the form

$$\Phi(P_k, 2\mathbf{i}\sinh u) = C_1(u)(\mathbf{i}e^u)^k + C_2(u)(\mathbf{i}e^u)^{-k},$$

for some $C_1(u)$ and $C_2(u)$ which can be determined by solving the system of equations

$$\begin{cases} C_1(u) + C_2(u) = \Phi(P_0, 2\mathbf{i}\sinh u) = 1 \\ C_1(u)\mathbf{i}e^u + C_2(u)(\mathbf{i}e^u)^{-1} = \Phi(P_1, 2\mathbf{i}\sinh u) = 2\mathbf{i}\sinh u. \end{cases}$$

After some calculations, we get $C_1(u) = e^{2u}/(1+e^{2u}), C_2(u) = 1/(1+e^{2u})$ and

$$\Phi(P_k, 2\mathbf{i}\sinh u) = \frac{e^{2u}}{1+e^{2u}}(\mathbf{i}e^u)^k + \frac{1}{1+e^{2u}}(\mathbf{i}e^u)^{-k}. \tag{6.7}$$

Hence

$$\begin{aligned}\Phi(C_k, 2\mathbf{i}\sinh u) &= \frac{e^{2u}}{1+e^{2u}}((\mathbf{i}e^u)^k - (\mathbf{i}e^u)^{k-2}) + \frac{1}{1+e^{2u}}((\mathbf{i}e^u)^{-k} - (\mathbf{i}e^u)^{-k+2}) - 2 \\ &= (\mathbf{i}e^u)^k + (\mathbf{i}e^u)^{-k} - 2 \\ &= (\mathbf{i}e^u)^k + (-\mathbf{i}e^{-u})^k - 2, \end{aligned} \tag{6.8}$$

and

$$\begin{aligned}&\Phi(P_k^{k-1}, 2\mathbf{i}\sinh u) \\ &= \mathbf{i}(e^u - e^{-u})((\mathbf{i}e^u)^{k-1} + (-\mathbf{i}e^{-u})^{k-1} - 2) - \frac{e^{2u}}{1+e^{2u}}(\mathbf{i}e^u)^{k-2} - \frac{1}{1+e^{2u}}(\mathbf{i}e^u)^{-k+2} \\ &= \frac{e^{2u}+1-e^{-2u}}{1+e^{2u}}(\mathbf{i}e^u)^k - \frac{e^{4u}-e^{2u}-1}{1+e^{2u}}(-\mathbf{i}e^{-u})^k - 2\mathbf{i}(e^u - e^{-u}).\end{aligned}$$

The characteristic polynomial of the linear recurrence relation (6.3) is the same as (6.4), and consequently has the two roots given in (6.5) and (6.6). Hence, the explicit expression of $\Phi(P_n^k, 2\mathbf{i}\sinh u)$ is of the form

$$\Phi(P_n^k, 2\mathbf{i}\sinh u) = A(u)(\mathbf{i}e^u)^n + B(u)(\mathbf{i}e^u)^{-n},$$

where $A(u)$ and $B(u)$ are such that

$$\begin{cases} A(u)(\mathbf{i}e^u)^k + B(u)(\mathbf{i}e^u)^{-k} = \Phi(P_k^k, 2\mathbf{i}\sinh u) = (\mathbf{i}e^u)^k + (\mathbf{i}e^u)^{-k} - 2, \\ A(u)(\mathbf{i}e^u)^{k+1} + B(u)(\mathbf{i}e^u)^{-k-1} = \Phi(P_{k+1}^k, 2\mathbf{i}\sinh u) \\ \qquad = \frac{e^{2u}+1-e^{-2u}}{1+e^{2u}}(\mathbf{i}e^u)^{k+1} - \frac{e^{4u}-e^{2u}-1}{1+e^{2u}}(-\mathbf{i}e^{-u})^{k+1} - 2\mathbf{i}(e^u - e^{-u}). \end{cases}$$

Solving the system of equations we get

$$A(u) = \frac{e^{2u}(e^{2u}+2)}{(1+e^{2u})^2} + \frac{2(\mathbf{i}e^u)^{2-k}}{1+e^{2u}} + \frac{(\mathbf{i}e^u)^{2(2-k)}}{(1+e^{2u})^2}$$

and

$$B(u) = \frac{(\mathbf{i}e^u)^{2k}}{(1+e^{2u})^2} - \frac{2(\mathbf{i}e^u)^k}{1+e^{2u}} + \frac{2e^{2u}+1}{(1+e^{2u})^2}.$$

Therefore we obtain

$$\begin{aligned}\Phi(P_n^k, 2\mathbf{i}\sinh u) &= \left(\frac{e^{2u}(e^{2u}+2)}{(1+e^{2u})^2} + \frac{2(\mathbf{i}e^u)^{2-k}}{1+e^{2u}} + \frac{(\mathbf{i}e^u)^{2(2-k)}}{(1+e^{2u})^2}\right)(\mathbf{i}e^u)^n \\ &\quad + \left(\frac{(\mathbf{i}e^u)^{2k}}{(1+e^{2u})^2} - \frac{2(\mathbf{i}e^u)^k}{1+e^{2u}} + \frac{2e^{2u}+1}{(1+e^{2u})^2}\right)(\mathbf{i}e^u)^{-n} \qquad (6.9)\end{aligned}$$

which implies

$$\begin{aligned}|\Phi(P_n^k, 2\mathbf{i}\sinh u)| = e^{nu}\Bigg| &\frac{e^{2u}(e^{2u}+2)}{(1+e^{2u})^2} + \frac{2(\mathbf{i}e^u)^{2-k}}{1+e^{2u}} + \frac{(\mathbf{i}e^u)^{2(2-k)}}{(1+e^{2u})^2} + \frac{(\mathbf{i}e^u)^{2k-2n}}{(1+e^{2u})^2} \\ &- \frac{2(\mathbf{i}e^u)^{k-2n}}{1+e^{2u}} + \frac{(2e^{2u}+1)(\mathbf{i}e^u)^{-2n}}{(1+e^{2u})^2}\Bigg|.\end{aligned}$$

Since

$$\begin{aligned}&\int_0^{+\infty} \log\left(\frac{e^{nu}}{(2\sinh u)^n}\right)\cosh u\, du \\ &\qquad = -n\int_0^{+\infty} \log(1-e^{-2u})d(\sinh u) \\ &\qquad = -n\left(\left[\log(1-e^{-2u})\sinh u\right]_0^{+\infty} - \int_0^{+\infty}(\sinh u)d(\log(1-e^{-2u}))\right) \\ &\qquad = n\int_0^{+\infty} e^{-u}du = n,\end{aligned}$$

we end up with

$$\mathrm{En}(P_n^k) = \frac{4n}{\pi} + \frac{4}{\pi}\int_0^{+\infty} \log(Q_n^k(u))\cosh u\ \ du, \qquad (6.10)$$

where

$$\begin{aligned}Q_n^k(u) &= \Bigg|\frac{e^{2u}(e^{2u}+2)}{(1+e^{2u})^2} + \frac{2(\mathbf{i}e^u)^{2-k}}{1+e^{2u}} + \frac{(\mathbf{i}e^u)^{2(2-k)}}{(1+e^{2u})^2} + \frac{(\mathbf{i}e^u)^{2k-2n}}{(1+e^{2u})^2} \\ &\qquad - \frac{2(\mathbf{i}e^u)^{k-2n}}{1+e^{2u}} + \frac{(2e^{2u}+1)(\mathbf{i}e^u)^{-2n}}{(1+e^{2u})^2}\Bigg| \\ &= \Bigg|\frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{(\mathbf{i}e^u)^{-k} - (\mathbf{i}e^u)^{k-2(n+1)}}{e^{2u}+1} \\ &\qquad + \frac{(\mathbf{i}e^u)^{-2k+4} + (\mathbf{i}e^u)^{2k-2n}}{(e^{2u}+1)^2} + \frac{2e^{2u}+1}{(e^{2u}+1)^2}(\mathbf{i}e^u)^{-2n}\Bigg|. \qquad (6.11)\end{aligned}$$

If we fix k and let $n \to \infty$, we obtain the following result:

Proposition 6.8 For any fixed $k \geq 3$, we have

$$\lim_{n\to\infty}\left(\mathrm{En}(P_n^k) - \frac{4n}{\pi}\right) = C(k)$$
$$:= \frac{4}{\pi}\int_0^{+\infty} \log\left|\frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{(\mathbf{i}e^u)^{-k}}{e^{2u}+1} + \frac{(-e^{2u})^{2-k}}{(e^{2u}+1)^2}\right| \cosh u \ du.$$

An analogous result holds if we fix $\ell = n-k$:

Proposition 6.9 For any fixed $\ell \geq 0$, we have

$$\lim_{n\to\infty}\left(\mathrm{En}(P_n^{n-\ell}) - \frac{4n}{\pi}\right) = D(\ell)$$
$$:= \frac{4}{\pi}\int_0^{+\infty} \log\left|\frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} + \frac{(-e^{2u})^{-\ell}}{(e^{2u}+1)^2}\right| \cosh u \ du.$$

Finally, we can let k and $n-k$ tend to ∞ simultaneously:

Proposition 6.10 If both k and $n-k$ go to infinity, then

$$\mathrm{En}(P_n^k) - \frac{4n}{\pi} \to \frac{4}{\pi}\int_0^{+\infty} \cosh u \log \frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2}\, du$$
$$= \frac{6\sqrt{2}}{\pi}\arctan\sqrt{2} + \frac{4}{\pi} - 4 \approx -0.146499.$$

For example, $C(3) \approx -0.037$, $C(4) \approx -0.866$, $C(5) \approx -0.084$, $C(6) \approx 0.118$, $D(0) = 0$, $D(1) \approx -0.246$, $D(2) \approx -0.087$, $D(3) \approx -0.200$. Figure 6.3 shows plots of more values of $C(k)$ and $D(l)$.

In the following, we will use again the abbreviation defined in (4.18):

$$g_{a,n,r}(j) = a^j + (-1)^r a^{n-j}.$$

Remember that for all non-negative integers n, r and $a \in (0,1)$, the function $g_{a,n,r}$ is positive and decreasing on $[0, n/2)$. If r is even, then we have $g_{a,n,r}(j) = g_{a,n,r}(n-j)$ and hence $g_{a,n,r}$ is increasing for $j \in (\frac{n}{2}, n]$. But $g_{a,n,r}$ is decreasing for $j \in (\frac{n}{2}, n]$ if r is odd, since we then have $g_{a,n,r}(j) = -g_{a,n,r}(n-j)$.

6.2.2 The behavior of $\mathrm{En}(P_n^k)$ for different values of k

In this section, we are interested in the behavior of the energy of P_n^k as k varies. To this end, we need to distinguish two cases depending on the parity of n as well as three cases for the residue class of k

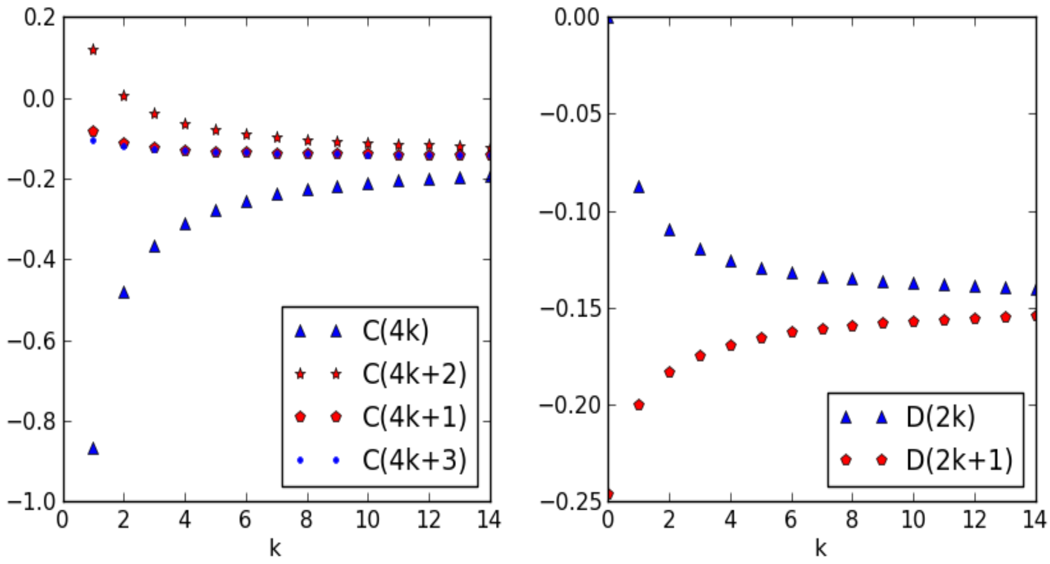


Figure 6.3: Plot of the functions $C(k)$ and $D(l)$

modulo 4. It is convenient to define a new variable $z = e^u$, for the sake of simplification.
Case 1: n is an odd integer.

1. $k \equiv 0 \mod 4$: From (6.11) we have

$$\begin{aligned} Q_n^k(u) &= \frac{z^2(z^2+2)}{(z^2+1)^2} - \frac{2z^2}{z^2+1}\left(z^{-k} - z^{-(2n+2-k)}\right) \\ &\qquad + \frac{z^{-2(k-2)} - z^{-2(n-k)}}{(z^2+1)^2} - \frac{2z^2+1}{(z^2+1)^2} z^{-2n} \\ &< \frac{z^4+2z^2+1}{(z^2+1)^2} + \frac{z^{-2(k-2)}-1}{(z^2+1)^2} < 1 \end{aligned}$$

and consequently

$$\mathrm{En}(P_n^k) < \frac{4n}{\pi}.$$

As we will see, this implies that no P_n^k, for $k \equiv 0 \mod 4$ and odd n, can be a candidate for the maximum energy in $\mathbb{U}_n$ because $\mathrm{En}(P_n^6) > \frac{4n}{\pi}$ (see inequalities (6.17) and (6.18)). They cannot be candidates for the second-largest energy in $\mathbb{U}_n$ either, because $\mathrm{En}(D_n) > \frac{4n}{\pi}$ for $n \geq 23$ (see inequalities (6.19) and (6.20)).

2. $k \equiv 2 \mod 4$:

$$\begin{aligned} Q_n^k(u) &= \left| \frac{z^2(z^2+2)}{(z^2+1)^2} + 2z^2 \frac{z^{-k} - z^{-(2n+2-k)}}{z^2+1} \right. \\ &\qquad \left. + \frac{z^{-2(k-2)} - z^{-2(n-k)}}{(z^2+1)^2} - \frac{2z^2+1}{(z^2+1)^2} z^{-2n} \right| \\ &= \frac{z^2(z^2+2)}{(z^2+1)^2} + \frac{2z^2}{z^2+1} g_{z^{-1},2(n+1),1}(k) \\ &\qquad + \frac{z^4}{(z^2+1)^2} g_{z^{-2},n+2,1}(k) - \frac{2z^2+1}{(z^2+1)^2} z^{-2n}. \end{aligned} \tag{6.12}$$

Equation (6.12) shows that $Q_n^k(u)$ is decreasing as a function of k in $[0, n]$ just as $g_{z^{-1},2(n+1),1}(k)$ and $g_{z^{-2},n+2,1}(k)$ are. Therefore we conclude that for all integers $n \geq k > 10$ (where $k \equiv 2 \mod 4$) and all real $u > 0$ we have $Q_n^6(u) > Q_n^{10}(u) > Q_n^k(u)$, which implies that

$$\mathrm{En}(P_n^6) > \mathrm{En}(P_n^{10}) > \mathrm{En}(P_n^k).$$

3. k is odd ($k \equiv 1 \mod 4$ or $k \equiv 3 \mod 4$): For this case, what we obtain from (6.11) is

$$\begin{aligned} (Q_n^k(u))^2 &= \left| \frac{z^2(z^2+2)}{(z^2+1)^2} \pm 2i \frac{z^{-(k-2)} + z^{-(2n-k)}}{z^2+1} \right. \\ &\qquad \left. - \frac{z^{-2(k-2)} - z^{-2(n-k)}}{(z^2+1)^2} - \frac{2z^2+1}{(z^2+1)^2} z^{-2n} \right|^2 \\ &= \left(\frac{z^2(z^2+2)}{(z^2+1)^2} - \frac{z^{-2(k-2)} - z^{-2(n-k)}}{(z^2+1)^2} - \frac{2z^2+1}{(z^2+1)^2} z^{-2n} \right)^2 \\ &\qquad + \left(2 \frac{z^{-(k-2)} + z^{-(2n-k)}}{z^2+1} \right)^2. \end{aligned}$$

Consider the second derivative of $\log(Q_n^k(u))^2$ with respect to k, which is

$$\begin{aligned} &\frac{\partial^2}{\partial k^2} \log\left((Q_n^k(u))^2\right) \\ &= \frac{\log^2 z}{(z^2+1)^8 (Q_n^k(u))^4} (8z^{-2n+2k+12} + 48z^{-2n+2k+10} + 96z^{-2n+2k+8} \\ &+ 64z^{-2n+2k+6} + 48z^{-2n-2k+14} + 240z^{-2n-2k+12} + 400z^{-2n-2k+10} \\ &+ 160z^{-2n-2k+8} + 64z^{-2n-2k+6} + 64z^{-2n-4k+14} + 64z^{-2n-4k+12} \\ &+ 64z^{-2n-4k+10} + 16z^{-2n-6k+14} + 8z^{-2n-6k+12} + 64z^{12-2n} + 256z^{10-2n} \\ &+ 384z^{8-2n} + 256z^{6-2n} + 16z^{-4n+4k+8} + 64z^{-4n+4k+6} + 64z^{-4n+4k+4} \end{aligned}$$

$$
\begin{aligned}
&+16z^{-4n+2k+12}+112z^{-4n+2k+10}+304z^{-4n+2k+8}+336z^{-4n+2k+6}\\
&+240z^{-4n+2k+4}+240z^{-4n-2k+12}+336z^{-4n-2k+10}+304z^{-4n-2k+8}\\
&+112z^{-4n-2k+6}+16z^{-4n-2k+4}+64z^{-4n-4k+12}+64z^{-4n-4k+10}\\
&+16z^{-4n-4k+8}+128z^{12-4n}+512z^{10-4n}+960z^{8-4n}+512z^{6-4n}\\
&+128z^{4-4n}+8z^{-6n+6k+4}+16z^{-6n+6k+2}+64z^{-6n+4k+6}+64z^{-6n+4k+4}\\
&+64z^{-6n+4k+2}+64z^{-6n+2k+10}+160z^{-6n+2k+8}+400z^{-6n+2k+6}\\
&+240z^{-6n+2k+4}+48z^{-6n+2k+2}+64z^{-6n-2k+10}+96z^{-6n-2k+8}\\
&+48z^{-6n-2k+6}+8z^{-6n-2k+4}+256z^{10-6n}+384z^{8-6n}+256z^{6-6n}\\
&+64z^{4-6n}+16z^{-8n+6k+4}+16z^{-8n+6k+2}+8z^{6k-8n}+64z^{-8n+4k+4}\\
&+64z^{-8n+4k+2}+16z^{4k-8n}+64z^{-8n+2k+8}+128z^{-8n+2k+6}+112z^{-8n+2k+4}\\
&+48z^{-8n+2k+2}+8z^{2k-8n}+8z^{16-2k}+48z^{14-2k}+112z^{12-2k}+128z^{10-2k}\\
&+64z^{8-2k}+16z^{16-4k}+64z^{14-4k}+64z^{12-4k}+8z^{16-6k}+16z^{14-6k}\\
&+16z^{12-6k})>0.
\end{aligned}
$$

This means that $\mathrm{En}(P_n^k)$ is convex as a function of k in this case (by differentiation under the integral sign). This shows that the maximum occurs at one of the ends.

Case 2: n is an even integer.

1. For $k \equiv 0 \mod 4$, we get

$$
\begin{aligned}
Q_n^k(u) &= \frac{z^2(z^2+2)}{(z^2+1)^2} - 2z^2\frac{z^{-k}+z^{-(2(n+1)-k)}}{z^2+1}\\
&\qquad + z^4\frac{z^{-2k}+z^{-2(n+2-k)}}{(z^2+1)^2} + \frac{2z^2+1}{(z^2+1)^2}z^{-2n}\\
&= \frac{z^2(z^2+2)}{(z^2+1)^2} - 2(z^4+z^2)\frac{z^{-k}+z^{-(2(n+1)-k)}}{(z^2+1)^2}\\
&\qquad + z^4\frac{z^{-2k}+z^{-2(n+2-k)}}{(z^2+1)^2} + \frac{2z^2+1}{(z^2+1)^2}z^{-2n}\\
&< \frac{z^2(z^2+2)}{(z^2+1)^2} - 2(z^4+z^2)\frac{z^{-(2(n+1)-k)}}{(z^2+1)^2} + \frac{z^{-2(n-k)}}{(z^2+1)^2}\\
&< \frac{z^2(z^2+2)+1}{(z^2+1)^2} = 1.
\end{aligned}
$$

Exactly as in the corresponding subcase for odd n, this implies

$$\mathrm{En}(P_n^k) < \frac{4n}{\pi}$$

for all integers $n \geq k > 0$ with $k \equiv 0 \mod 4$. Therefore P_n^k cannot be the unicyclic graph with largest or second-largest energy in this case (see again inequalities (6.17), (6.18), (6.19) and (6.20)).

2. For $k \equiv 2 \mod 4$, equation (6.11) gives

$$Q_n^k(u) = \frac{z^2(z^2+2)}{(z^2+1)^2} + 2z^2\frac{z^{-k}+z^{-(2(n+1)-k)}}{z^2+1} + \frac{z^{-2(k-2)}+z^{-2(n-k)}}{(z^2+1)^2} + \frac{2z^2+1}{(z^2+1)^2}z^{-2n}.$$

Similar to the case of odd n and k, we obtain

$$\begin{aligned}
&\frac{\partial^2}{\partial k^2}\log\left(Q_n^k(u)\right)\\
&= \frac{\log^2 z}{(z^2+1)^4(Q_n^k(u))^4}\big(4z^{-2n+2k+4} + 8z^{-2n+2k+2} + 2z^{-2n+k+6} + 24z^{-2n+k+4}\\
&+ 22z^{-2n+k+2} + 22z^{-2n-k+6} + 24z^{-2n-k+4} + 2z^{-2n-k+2} + 8z^{-2n-2k+6}\\
&+ 4z^{-2n-2k+4} + 16z^{6-2n} + 48z^{4-2n} + 16z^{2-2n} + 2z^{-4n+3k+2} + 2z^{3k-4n}\\
&+ 8z^{-4n+2k+2} + 4z^{2k-4n} + 4z^{-4n+k+4} + 6z^{-4n+k+2} + 2z^{k-4n} + 2z^{8-k}\\
&+ 6z^{6-k} + 4z^{4-k} + 4z^{8-2k} + 8z^{6-2k} + 2z^{8-3k} + 2z^{6-3k}\big) > 0,
\end{aligned}$$

which means that $\mathrm{En}(P_n^k)$ is convex as a function of k. Therefore for all integers $k \equiv 2 \mod 4$ such that $n-2 > k > 6$ we have

$$\begin{cases}\max\{\mathrm{En}(P_n^6), \mathrm{En}(C_n)\} > \mathrm{En}(P_n^k) & \text{for } n \equiv 2 \mod 4,\\ \max\{\mathrm{En}(P_n^6), \mathrm{En}(P_n^{n-2})\} > \mathrm{En}(P_n^k) & \text{for } n \equiv 0 \mod 4.\end{cases}$$

3. For $k \equiv 3 \mod 4$ or $k \equiv 1 \mod 4$ we get (from (6.11))

$$\begin{aligned}
(Q_n^k(u))^2 &= \left|\frac{z^2(z^2+2)}{(z^2+1)^2} \pm 2iz^2\frac{z^{-k}-z^{-(2n+2-k)}}{z^2+1} - \frac{z^{-2(k-2)}+z^{-2(n-k)}}{(z^2+1)^2} + \frac{2z^2+1}{(z^2+1)^2}z^{-2n}\right|^2\\
&= \left(\frac{z^2(z^2+2)}{(z^2+1)^2} - z^4\frac{g_{z^{-2},n+2,2}(k)}{(z^2+1)^2} + \frac{2z^2+1}{(z^2+1)^2}z^{-2n}\right)^2\\
&\quad + \left(2z^2\frac{g_{z^{-1},2n+2,1}(k)}{z^2+1}\right)^2. \qquad (6.13)
\end{aligned}$$

We can see from (6.13) that $(Q_n^k(u))^2$ (and hence $Q_n^k(u)$) decreases as a function of k on $[\frac{n+2}{2}, n]$: this is because the two functions $-g_{z^{-2},n+2,2}$ and $g_{z^{-1},2n+2,1}$ both decrease on this interval. Furthermore, let

$$\begin{aligned}
B(z,n) &= z^2(z^2+2) + (2z^2+1)z^{-2n}\\
&= (z^2+1)^2 + (z^2+1)^2z^{-2n} - 1 - z^{-2(n-2)}\\
&= (z^2+1)^2(1+z^{-2n}) - 1 - z^{-2(n-2)}
\end{aligned}$$

so that we have

$$
\begin{aligned}
(Q_n^k(u))^2 &= \left(\frac{B(z,n)}{(z^2+1)^2} - z^4\frac{g_{z^{-2},n+2,2}(k)}{(z^2+1)^2}\right)^2 + \left(2z^2\frac{g_{z^{-1},2n+2,1}(k)}{z^2+1}\right)^2 \\
&= \frac{B^2(z,n)}{(z^2+1)^4} - 2B(z,n)z^4\frac{g_{z^{-2},n+2,2}(k)}{(z^2+1)^4} \\
&\qquad + z^8\frac{g^2_{z^{-2},n+2,2}(k)}{(z^2+1)^4} + 4z^4\frac{g^2_{z^{-1},2n+2,1}(k)}{(z^2+1)^2},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial}{\partial k}(Q_n^k(u))^2 &= \frac{g_{z^{-2},n+2,1}(k)}{(z^2+1)^4}4B(z,n)z^4\log z \\
&- \frac{g_{z^{-2},n+2,2}(k)g_{z^{-2},n+2,1}(k)}{(z^2+1)^4}4z^8\log z - \frac{g_{z^{-},2n+2,1}(k)g_{z^{-},2n+2,2}(k)}{(z^2+1)^2}8z^4\log z \\
&= \frac{g_{z^{-2},n+2,1}(k)}{(z^2+1)^4}4B(z,n)z^4\log z \\
&\qquad - \frac{g_{z^{-4},n+2,1}(k)}{(z^2+1)^4}4z^8\log z - \frac{g_{z^{-2},2n+2,1}(k)}{(z^2+1)^2}8z^4\log z \\
&\le \frac{4z^4\log z}{(z^2+1)^2}\Bigg((1+z^{-2n})g_{z^{-2},n+2,1}(k) - \frac{g_{z^{-2},n+2,1}(k)}{(z^2+1)^2} \\
&\qquad - z^4\frac{g_{z^{-4},n+2,1}(k)}{(z^2+1)^2} - g_{z^{-2},2n+2,1}(k)\Bigg) \text{ for } k \le \frac{n+2}{2} \\
&= \frac{4z^4\log z}{(z^2+1)^2}\Big(-z^{-2(n+2-k)} + z^{-2(n+k)} - z^{-2(2n+2-k)} \\
&\qquad - \frac{g_{z^{-2},n+2,1}(k)}{(z^2+1)^2} - z^4\frac{g_{z^{-4},n+2,1}(k)}{2(z^2+1)^2} + z^{-2(2n+2-k)}\Big) \\
&= \frac{4z^4\log z}{(z^2+1)^2}\Bigg(-z^{-2(n+2-k)} + z^{-2(n+k)} - \frac{g_{z^{-2},n+2,1}(k)}{(z^2+1)^2} \\
&\qquad - z^4\frac{g_{z^{-4},n+2,1}(k)}{(z^2+1)^2}\Bigg) < 0 \text{ (remember that } k \ge 3).
\end{aligned}
$$

This means that $Q_n^k(u)$ is also decreasing on the interval $[1, \frac{n+2}{2}]$ and thus on the entire interval $[1, n]$. Therefore for all odd k such that $3 < k \le n$ and all $u > 0$ we have

$$Q_n^3(u) > Q_n^k(u)$$

which implies

$$\mathrm{En}(P_n^3) > \mathrm{En}(P_n^k). \tag{6.14}$$

We can conclude now that the tadpole with largest energy is an element of

$$
\begin{cases}
\{P_n^3, P_n^6, P_n^{n-2}\} & \text{if } n \equiv 0 \mod 4, \\
\{P_n^3, P_n^6, C_n\} & \text{otherwise.}
\end{cases}
\tag{6.15}
$$

Once we will be able to see that $\mathrm{En}(P_n^6)$ is the n-vertex tadpole with largest energy, then it follows that the tadpole with second-largest energy must be an element of

$$\begin{cases} \{P_n^3, P_n^{10}, P_n^{n-2}\} \text{ if } n \equiv 0 \mod 4, \\ \{P_n^3, P_n^{10}, C_n\} \quad \text{otherwise.} \end{cases} \tag{6.16}$$

6.2.3 Estimating the energy in special cases

We now collect estimates for the energy of the remaining graphs to be considered. For the two graphs P_n^6 and D_n, which are conjectured to have large energy, we provide estimates from below. On the other hand, we determine upper estimates for $\mathrm{En}(C_n)$, $\mathrm{En}(P_n^3)$, $\mathrm{En}(P_n^{10})$ and $\mathrm{En}(P_n^{n-2})$. The main theorems of this chapter are then obtained by combining these estimates. Most of our estimates are obtained from the integral formula (6.10). It simplifies formulas to use the substitution $y = e^{-u}$.

For even n and $k = 6$, (6.11) gives

$$\begin{aligned} &Q_n^6(u) \\ &= \left| \frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{-e^{-6u} - e^{(4-2n)u}}{e^{2u}+1} + \frac{e^{-8u} + e^{(12-2n)u}}{(e^{2u}+1)^2} + \frac{2e^{2u}+1}{(e^{2u}+1)^2}e^{-2nu} \right| \\ &= \frac{y^{-4} + 2y^{-2} + y^8 + y^{2n-12} + 2y^{2n-2} + y^{2n}}{(y^{-2}+1)^2} + \frac{2y^4 + y^{2n-6}}{y^{-2}+1}, \\ &\geq \frac{y^{-4} + 2y^{-2} + y^8}{(y^{-2}+1)^2} + \frac{2y^4}{y^{-2}+1}, \end{aligned}$$

and hence (6.10) gives

$$\begin{aligned} \mathrm{En}(P_n^6) &> \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 \frac{y^2+1}{y^2} \log\left(\frac{y^{-4} + 2y^{-2} + y^8}{(y^{-2}+1)^2} + \frac{2y^4}{y^{-2}+1} \right) dy \\ &> \frac{4n}{\pi} + \frac{0.370}{\pi}. \end{aligned} \tag{6.17}$$

For odd $n \geq 17$, we obtain

$$\begin{aligned} &Q_n^6(u) \\ &= \left| \frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{-e^{-6u} + e^{(4-2n)u}}{e^{2u}+1} + \frac{e^{-8u} - e^{(12-2n)u}}{(e^{2u}+1)^2} - \frac{2e^{2u}+1}{(e^{2u}+1)^2}e^{-2nu} \right| \\ &= \frac{y^{-4} + 2y^{-2} + y^8 - y^{2n-12} - 2y^{2n-2} - y^{2n}}{(y^{-2}+1)^2} + \frac{2y^4 - y^{2n-6}}{y^{-2}+1}, \\ &\geq \frac{y^{-4} + 2y^{-2} + y^8 - y^{2\cdot17-12} - 2y^{2\cdot17-2} - y^{2\cdot17}}{(y^{-2}+1)^2} + \frac{2y^4 - y^{2\cdot17-6}}{y^{-2}+1}, \end{aligned}$$

and

$$\begin{aligned}\mathrm{En}(P_n^6) &\geq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 \frac{y^2+1}{y^2}\log\left(\frac{y^{-4}+2y^{-2}+y^8-y^{2\cdot 17-12}-2y^{2\cdot 17-2}-y^{2\cdot 17}}{(y^{-2}+1)^2}\right.\\ &\qquad\left.+\frac{2y^4-y^{2\cdot 17-6}}{y^{-2}+1}\right)dy\\ &> \frac{4n}{\pi}+\frac{0.226}{\pi}. \qquad (6.18)\end{aligned}$$

In a similar way as (6.10) was obtained, we also get

$$\mathrm{En}(D_n) = \frac{4n}{\pi} + \frac{4}{\pi}\int_0^{+\infty}\log\left|\frac{\Phi(D_n, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\right|\cosh u\ du.$$

Using Lemma 3.8 we have

$$\Phi(D_n, x) = (x^2-1)\Phi(P_{n-2}^6, x) - x\Phi(C_6, x)\Phi(P_{n-9}, x),$$

and hence

$$\begin{aligned}\Phi(D_n, 2\mathbf{i}\sinh u) &= (-e^{2u}-e^{-2u}+1)\Phi(P_{n-2}^6, 2\mathbf{i}\sinh u)\\ &\qquad - \mathbf{i}(e^u-e^{-u})\Phi(C_6, 2\mathbf{i}\sinh u)\Phi(P_{n-9}, 2\mathbf{i}\sinh u).\end{aligned}$$

Since (see (6.9))

$$\begin{aligned}&(-e^{2u}-e^{-2u}+1)\frac{\Phi(P_{n-2}^6, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\\ &= \frac{(-e^{2u}-e^{-2u}+1)}{(\mathbf{i}e^u)^2}\left(\frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{(\mathbf{i}e^u)^{-6}-(\mathbf{i}e^u)^{6-2(n-2+1)}}{e^{2u}+1}\right.\\ &\qquad\left.+\frac{(\mathbf{i}e^u)^{-2\cdot 6+4}+(\mathbf{i}e^u)^{2\cdot 6-2(n-2)}}{(e^{2u}+1)^2}+\frac{2e^{2u}+1}{(e^{2u}+1)^2}(\mathbf{i}e^u)^{-2(n-2)}\right)\\ &= (1+y^4-y^2)\left(\frac{y^{-4}+2y^{-2}+(-1)^n(2y^{2n-6}+y^{2n-4})}{(1+y^{-2})^2}\right.\\ &\qquad\left.+2\frac{y^4+(-1)^n y^{2n-10}}{1+y^{-2}}+\frac{y^8+(-1)^n y^{2n-16}}{(1+y^{-2})^2}\right)\\ &= \frac{(1-y^2+y^4)^2}{1+y^2}\left(1+2y^2+y^6+(-1)^n(y^{2n-6}+2y^{2n-8}+y^{2n-12})\right)\end{aligned}$$

and (see (6.7) and (6.8))

$$\begin{aligned}&\mathbf{i}(e^u-e^{-u})\Phi(C_6, 2\mathbf{i}\sinh u)\frac{\Phi(P_{n-9}, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\\ &\quad= \mathbf{i}(y^{-1}-y)\left((\mathbf{i}y^{-1})^6+(-\mathbf{i}y)^6-2\right)\frac{\frac{1}{1+y^2}(\mathbf{i}y^{-1})^{n-9}+\frac{y^2}{1+y^2}(\mathbf{i}y^{-1})^{-(n-9)}}{(\mathbf{i}y^{-1})^n}\\ &\quad= (y^4-1)(y^4-y^2+1)^2\left(y^2-(-1)^n y^{2n-14}\right),\end{aligned}$$

we obtain

$$\left|\frac{\Phi(D_n, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\right| = \frac{(1-y^2+y^4)^2}{1+y^2}\big(1+3y^2+y^4-y^8 + (-1)^n(y^{2n-6}+3y^{2n-8}+y^{2n-10}-y^{2n-14})\big).$$

- For even n, we now have

$$\left|\frac{\Phi(D_n, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\right| \geq t_1(y,n) := \frac{(1-y^2+y^4)^2}{1+y^2}\left(1+3y^2+y^4-y^8-y^{2n-14}\right)$$

  and thus

$$\begin{aligned}\mathrm{En}(D_n) &\geq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 (y^{-2}+1)\log t_1(y,28)dy \text{ for all even } n \geq 28\\ &> \frac{4n}{\pi} + \frac{0.168}{\pi}. \end{aligned} \tag{6.19}$$

- For odd n we have

$$\left|\frac{\Phi(D_n, 2\mathbf{i}\sinh u)}{(\mathbf{i}e^u)^n}\right| \geq t_2(y,n) := \frac{(1-y^2+y^4)^2}{1+y^2}\left(1+3y^2+y^4-y^8-y^{2n-6}-3y^{2n-8}\right)$$

  and thus

$$\begin{aligned}\mathrm{En}(D_n) &\geq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 (y^{-2}+1)\log t_2(y,29)dy \text{ for all odd } n \geq 29\\ &> \frac{4n}{\pi} + \frac{0.062}{\pi}. \end{aligned} \tag{6.20}$$

For the cycle C_n, explicit formulas are available:

$$\mathrm{En}(C_n) = \begin{cases} 4\cot\frac{\pi}{n} & n \equiv 0 \mod 4,\\ 4\csc\frac{\pi}{n} & n \equiv 2 \mod 4,\\ 2\csc\frac{\pi}{2n} & n \text{ odd.}\end{cases}$$

Detailed proof of this is provided in [1]. This gives us a trivial lower bound for odd $n \geq 3$:

$$\mathrm{En}(C_n) = 2\frac{1}{\sin\frac{\pi}{2n}} > 2\frac{1}{\frac{\pi}{2n}} = \frac{4n}{\pi}. \tag{6.21}$$

For an upper bound, we notice that $\cot x < \csc x$, and that the function $\csc x - 1/x$ is increasing for $x < \pi$. Hence we have

$$\operatorname{En}(C_n) \leq 4\csc\frac{\pi}{n} \leq \frac{4n}{\pi} + 4\csc\frac{\pi}{40} - \frac{160}{\pi} < \frac{4n}{\pi} + \frac{0.165}{\pi} < \operatorname{En}(D_n) \tag{6.22}$$

for even $n \geq 40$, and the inequality $\operatorname{En}(C_n) < \operatorname{En}(D_n)$ can also be verified directly for even $n \in [28, 38]$. Likewise,

$$\operatorname{En}(C_n) = 2\csc\frac{\pi}{2n} \leq \frac{4n}{\pi} + 2\csc\frac{\pi}{58} - \frac{116}{\pi} < \frac{4n}{\pi} + \frac{0.057}{\pi} < \operatorname{En}(D_n) \tag{6.23}$$

for odd $n \geq 29$.

To estimate $\operatorname{En}(P_n^3)$, we first evaluate (6.11) for $k = 3$ and get

$$\begin{aligned}(Q_n^3(u))^2 = \Bigg| &\frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2\mathbf{i}e^{2u}\frac{e^{-3u} - (-1)^n e^{(1-2n)u}}{e^{2u}+1} \\ &- \frac{e^{-2u} + (-1)^n e^{(6-2n)u}}{(e^{2u}+1)^2} + \frac{2e^{2u}+1}{(e^{2u}+1)^2}(\mathbf{i}e^u)^{-2n}\Bigg|^2 \\ = &\left(\frac{y^{-4} + 2y^{-2} - y^2 - (-1)^n y^{2n-6} + (-1)^n(2y^{2n-2} + y^{2n})}{(y^{-2}+1)^2}\right)^2 \\ &+ \left(\frac{2y - 2(-1)^n y^{2n-3}}{y^{-2}+1}\right)^2.\end{aligned} \tag{6.24}$$

- If n is even, then we have

$$\begin{aligned}&(Q_n^3(u))^2 \\ &= \left(\frac{y^{-4} + 2y^{-2} - y^2 - y^{2n-6} + 2y^{2n-2} + y^{2n}}{(y^{-2}+1)^2}\right)^2 + \left(\frac{2y - 2y^{2n-3}}{y^{-2}+1}\right)^2 \\ &\leq u_1(y,n) := \left(\frac{1 + 2y^2 - y^6 + y^{2n+2} + y^{2n+4}}{(y^2+1)^2}\right)^2 + \left(\frac{2y^3}{y^2+1}\right)^2.\end{aligned}$$

Therefore

$$\begin{aligned}\operatorname{En}(P_n^3) &= \frac{4n}{\pi} + \frac{2}{\pi}\int_0^{+\infty} \log((Q_6^3(u))^2)\cosh u\,du \\ &< \frac{4n}{\pi} + \frac{1}{\pi}\int_0^1 (y^{-2}+1)\log u_1(y,28)dy \text{ for all even } n \geq 28 \\ &\leq \frac{4n}{\pi} - \frac{0.100}{\pi} < \operatorname{En}(D_n) \text{ (see (6.19))}\end{aligned} \tag{6.25}$$

and

$$\begin{aligned}\operatorname{En}(P_n^3) &< \frac{4n}{\pi} + \frac{1}{\pi}\int_0^1 (y^{-2}+1)\log u_1(y,6)dy \text{ for all even } n \geq 6 \\ &\leq \frac{4n}{\pi} - \frac{0.028}{\pi} < \operatorname{En}(P_n^6) \text{ (comparing with (6.17))}.\end{aligned} \tag{6.26}$$

- For odd n we obtain from (6.24) that

$$(Q_n^3(u))^2 = \left(\frac{y^{-4}+2y^{-2}-y^2+y^{2n-6}-2y^{2n-2}-y^{2n}}{(y^{-2}+1)^2}\right)^2 + \left(\frac{2y+2y^{2n-3}}{y^{-2}+1}\right)^2$$
$$\leq u_2(y,n) := \left(\frac{1+2y^2-y^6+y^{2n-2}}{(y^2+1)^2}\right)^2 + \left(\frac{2y^3+2y^{2n-1}}{y^2+1}\right)^2.$$

Hence, with use of (6.18) and (6.21), we obtain

$$\begin{aligned}\mathrm{En}(P_n^3) &\leq \frac{4n}{\pi} + \frac{1}{\pi}\int_0^1 (y^{-2}+1)\log u_2(y,17)dy \text{ for } n \geq 17 \\ &\leq \frac{4n}{\pi} < \min\{\mathrm{En}(C_n), \mathrm{En}(P_n^6)\}. \end{aligned} \tag{6.27}$$

Furthermore, we also have

$$\mathrm{En}(P_n^3) \leq \frac{4n}{\pi} < \mathrm{En}(D_n) \tag{6.28}$$

for $n \geq 23$. The last inequality is obtained in the same way as (6.20) replacing $t_2(y,29)$ by $t_2(y,23)$.

Taking $k=10$, (6.11) gives

$$\begin{aligned}Q_n^{10}(u) &= \left| \frac{e^{2u}(e^{2u}+2)}{(e^{2u}+1)^2} - 2e^{2u}\frac{(\mathbf{i}e^u)^{-10}-(\mathbf{i}e^u)^{10-2(n+1)}}{e^{2u}+1} \right. \\ &\qquad \left. + \frac{(\mathbf{i}e^u)^{-2\cdot 10+4}+(\mathbf{i}e^u)^{2\cdot 10-2n}}{(e^{2u}+1)^2} + \frac{2e^{2u}+1}{(e^{2u}+1)^2}(\mathbf{i}e^u)^{-2n} \right| \\ &= \frac{y^{-4}+2y^{-2}+(-1)^n(2y^{2n-2}+y^{2n})+y^{16}+(-1)^n y^{2n-20}}{(y^{-2}+1)^2} \\ &\qquad + 2\frac{y^8+(-1)^n y^{2n-10}}{y^{-2}+1}.\end{aligned}$$

- For even n we have

$$Q_n^{10}(u) = v_1(y,n) := \frac{1+2y^2+2y^{2n+2}+y^{2n+4}+y^{20}+y^{2n-16}}{(y^2+1)^2} + 2\frac{y^{10}+y^{2n-8}}{y^2+1}.$$

This and (6.19) lead to

$$\begin{aligned}\mathrm{En}(P_n^{10}) &\leq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 (y^{-2}+1)\log v_1(y,28)dy \text{ for } n \geq 28 \\ &\leq \frac{4n}{\pi} + \frac{0.092}{\pi} < \mathrm{En}(D_n). \end{aligned} \tag{6.29}$$

- For odd n we have

$$\begin{aligned} Q_n^{10}(u) &= \frac{1+2y^2-2y^{2n+2}-y^{2n+4}+y^{20}-y^{2n-16}}{(y^2+1)^2} + 2\frac{y^{10}-y^{2n-8}}{y^2+1} \\ &\leq v_2(y) := \frac{1+2y^2+y^{20}}{(y^2+1)^2} + \frac{2y^{10}}{y^2+1}. \end{aligned}$$

With (6.20), this implies

$$\begin{aligned} \mathrm{En}(P_n^{10}) &\leq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 (y^{-2}+1)\log v_2(y)dy \\ &\leq \frac{4n}{\pi} + \frac{0.016}{\pi} < \mathrm{En}(D_n). \end{aligned} \tag{6.30}$$

We only have to estimate $\mathrm{En}(P_n^{n-2})$ for $n \equiv 0 \mod 4$, in which case if we take $k = n-2$ (6.11) becomes

$$\begin{aligned} &Q_n^{n-2}(u) \\ &= \frac{e^{4u}+2e^{2u}+2e^{-2(n-1)u}+e^{-2nu}}{(e^{2u}+1)^2} + 2\frac{e^{-(n-4)u}+e^{-(n+2)u}}{e^{2u}+1} + \frac{e^{-2(n-4)u}+e^{-4u}}{(e^{2u}+1)^2} \\ &= w(y,n) := \frac{1+2y^2+2y^{2n+2}+y^{2n+4}+y^{2n-4}+y^8}{(y^2+1)^2} + 2\frac{y^{n-2}+y^{n+4}}{y^2+1}. \end{aligned}$$

This and (6.19) imply that for $n \geq 28$ we have

$$\begin{aligned} \mathrm{En}(P_n^{n-2}) &\leq \frac{4n}{\pi} + \frac{2}{\pi}\int_0^1 (y^{-2}+1)\log w(y,28)dy \\ &\leq \frac{4n}{\pi} - \frac{0.02}{\pi} < \mathrm{En}(D_n). \end{aligned} \tag{6.31}$$

6.3 Proofs of the main theorems

We are now ready to prove the three main results:

Proof of Theorem 6.7: Let B_n be an n-vertex non-bipartite unicyclic graph with maximum energy. By Theorem 6.2, we know that B_n has to be some P_n^k for some odd k. Furthermore, (6.15) implies that $B_n = P_n^3$ if n is even, and $B_n \in \{P_n^3, C_n\}$ if n is odd. Hence (6.27) completes the proof. □

Proof of Conjecture 6.1: Let U_n be an unicyclic graph with maximum energy. It follows from the two Theorems 6.3 and 6.7 that $U_n \in \{P_n^3, P_n^6, C_n\}$. The inequalities (6.22), (6.23), (6.26) and (6.27) show that P_n^6 always wins against P_n^3 and C_n, for even $n \geq 40$ and odd $n \geq 29$. Note that $\mathrm{En}(D_n) < \mathrm{En}(P_n^6)$ for all $n \geq 6$, see Theorem 6.3. The cases of small values of n can be checked directly. □

Proof of Theorem 6.5: From Theorem 6.6 and (6.16) we deduce that the n-vertex unicyclic graph with second largest energy, must be an element of

$$\begin{cases} \{P_n^3, P_n^{10}, D_n, P_n^{n-2}\} & \text{if } n \equiv 0 \mod 4, \\ \{P_n^3, P_n^{10}, D_n, C_n\} & \text{otherwise.} \end{cases}$$

(6.22), (6.23), (6.25), (6.28), (6.29), (6.30) and (6.31) show that D_n wins against any of the other candidates. $\square$

It is very likely that the same approach can also be used to characterize the unicyclic graph with third-largest, fourth-largest, ... energy, although the number of cases to be considered will become considerable. Propositions 6.8, 6.9 and 6.10 also show that there are lots of unicyclic graphs whose energy comes close (bounded difference) to the maximum value.

Part III

Number of subtrees, spectral moments, and greedy trees

Chapter 7

Introduction and preliminary observations

The so-called greedy trees have been shown to be extremal among trees with a given degree sequence with respect to many graph invariants such as the Wiener index (sum of all distances) [84, 92] and related distance-based invariants [75], the spectral radius [10] and Laplacian spectral radius [9, 90], etc. These trees are constructed from a given degree sequence by a simple greedy algorithm that assigns the highest degree to the root, the second-, third-, ... highest degrees to the root's neighbors, and so on – a formal definition will be given later in this chapter.

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a graph G. The k-th spectral moment of G is defined by

$$\mathrm{M}_k(G) = \sum_{k=0}^{n} \lambda_i^k. \tag{7.1}$$

In Chapter 8, we show that for any non-negative integer k the greedy tree has the maximum k-th spectral moment among all trees with the same degree sequence. It is also shown that if $D = (d_1, \ldots, d_n)$ and $B = (b_1, \ldots, b_n)$ are degree sequences of trees which satisfy

$$\sum_{i=1}^{l} b_i \leq \sum_{i=1}^{l} d_i \tag{7.2}$$

for all $1 \leq l \leq n$ (i.e., D majorizes B), then the greedy tree with degree sequence D has larger k-th spectral moment than the one with degree sequence B, cf. Theorem 5.28 in Chapter 5.

Similar results are obtained in Chapter 9 using number of subtrees instead of spectral moments. The greedy trees were shown in [94] to have the maximum number of subtrees among all trees with given degree sequence. The analogous problem for the minimum was studied recently in [93]. Further recent results on maxima and minima of the number of subtrees under various restrictions can be found in [66, 67]. In Chapter 9, the main result is the fact that the greedy tree not only maximizes the total number of subtrees but actually the number of subtrees of any given order. A similar result was achieved recently for distance-based graph invariants: in [75] it was shown that the number of pairs of vertices whose distance is at most k is maximized by the greedy tree (given the degree sequence) for every k. We also show in the same chapter that if we count only subtrees containing the root having a given number of leaves, then the maximum number is still obtained for the greedy tree. Additional result comparing the number of k-vertex subtrees of greedy trees with different degree sequences leads to many corollaries.

The following types of trees and forests will play main roles:

Definition 7.1 Let F be a rooted forest where the maximum height of any component is $k-1$. The *leveled degree sequence* of F is the sequence

$$D = (D_1, \ldots, D_k), \tag{7.3}$$

where, for any $1 \leq i \leq k$, D_i is the non-increasing sequence formed by the degrees of the vertices of F at the i^{th} level (i.e., vertices of distance $i-1$ from the root in the respective component).

For convenience, we denote the "number of levels" in D by $L(D)$ (maximum height plus one), evidently $L(D) = k$ in (7.3).

The greedy trees have been defined in various equivalent ways in previous works [10, 75, 84, 90]. For our purposes, we begin with the definitions of level greedy trees and forests. Note the labeling of vertices, because we will always use it for vertices of similar forests.

Definition 7.2 The *level greedy forest* with leveled degree sequence

$$D = ((i_{1,1}, \ldots, i_{1,k_1}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n})) \tag{7.4}$$

is obtained using the following "greedy algorithm":

(i) Label the vertices of the first level by $g_1^1, \ldots, g_{k_1}^1$, and assign degrees to these vertices such that $\deg g_j^1 = i_{1,j}$ for all j.

(ii) Assume that the vertices of the h^{th} level have been labeled $g_1^h, \ldots, g_{k_h}^h$ and a degree has been assigned to each of them. Then for all $1 \leq j \leq k_h$ label the neighbors of g_j^h at the $(h+1)^{\text{th}}$ level, if any, by

$$g^{h+1}_{1+\sum_{m=1}^{j-1}(i_{h,m}-1)}, \ldots, g^{h+1}_{\sum_{m=1}^{j}(i_{h,m}-1)},$$

and assign degrees to the newly labeled vertices such that $\deg g_j^{h+1} = i_{h+1,j}$ for all j.

The level greedy forest with leveled degree sequence D is denoted by $G(D)$ (Figure 7.1).

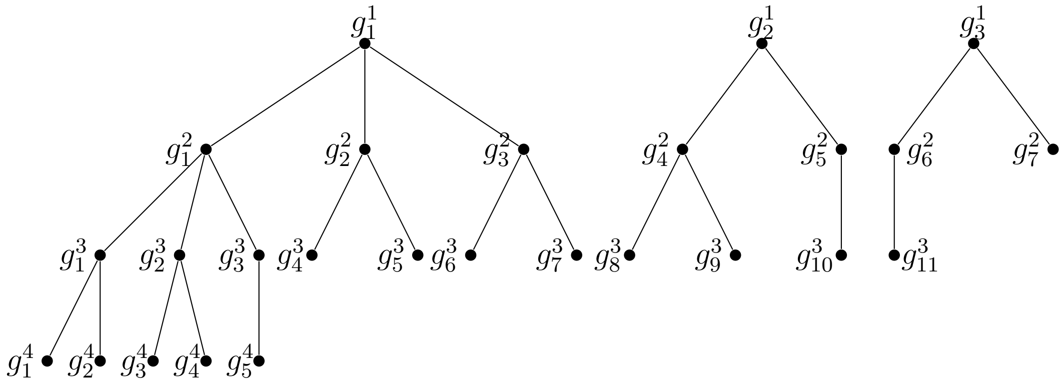


Figure 7.1: A level greedy forest

Definition 7.3 A connected level greedy forest is called a *level greedy tree*.

In analogy to (rooted) level greedy trees, we will also need an edge-rooted version:

Definition 7.4 The *edge-rooted* level greedy tree with leveled degree sequence

$$D = ((i_{1,1}, i_{1,2}), (i_{2,1}, \dots, i_{2,k_2}), \dots, (i_{n,1}, \dots, i_{n,k_n}))$$

is obtained from the two-component level greedy forest with leveled degree sequence

$$((i_{1,1} - 1, i_{1,2} - 1), (i_{2,1}, \dots, i_{2,k_2}), \dots, (i_{n,1}, \dots, i_{n,k_n}))$$

by joining the two roots.

Finally, we define greedy trees and greedy forests:

Definition 7.5 If a root in a tree can be chosen such that it becomes a level greedy tree whose leveled degree sequence, as given in (7.4), satisfies

$$\min\{i_{j,1}, \dots, i_{j,k_j}\} \geq \max\{i_{j+1,1}, \dots, i_{j+1,k_{j+1}}\}$$

for all $1 \leq j \leq n-1$, then it is called a *greedy tree* (Figure 7.2). In the case that D is a degree sequence (as opposed to a leveled degree sequence), we use $G(D)$ to denote the greedy tree with degree sequence D.

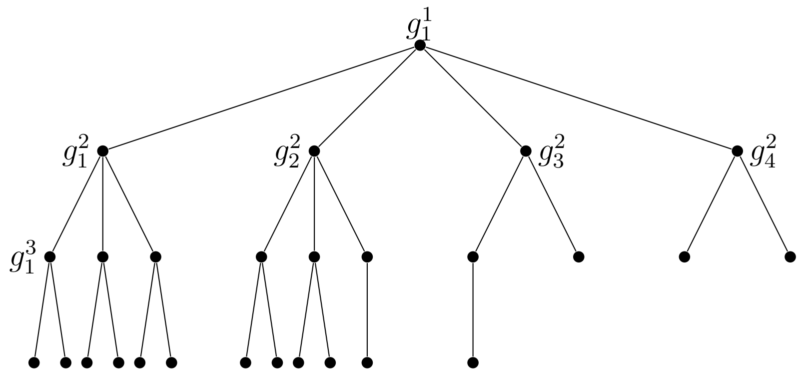


Figure 7.2: A greedy tree (only the labels of the first six vertices are shown)

Proofs of main theorems in the next two chapters use the following key remark, which uses an observation from [78]:

Remark 7.6 It is shown in [78] that, with a given degree sequence, a tree T which satisfies the following "semi-regular property" is a greedy tree: given any path with end vertices $u, v \in V(T)$, the set of subtrees $\{T_u^1, \ldots, T_u^a\}$ attached to u and the set of subtrees $\{T_v^1, \ldots, T_v^b\}$ attached to v (such that $v \notin T_u^i$ and $u \notin T_v^j$ for each i and j) satisfy

$$a \geq b \text{ and } \min\{|V(T_u^1)|, \ldots, |V(T_u^a)|\} \geq \max\{|V(T_v^1)|, \ldots, |V(T_v^b)|\}$$

or

$$b \geq a \text{ and } \min\{|V(T_v^1)|, \ldots, |V(T_v^b)|\} \geq \max\{|V(T_u^1)|, \ldots, |V(T_u^a)|\}.$$

Note that if a tree is level greedy with respect to any possible choice of vertex or edge as root, then it satisfies the "semi-regular property".

Definition 7.7 A forest with components $F_1, \ldots, F_t$ each of which is a greedy tree is called *greedy forest* if

$$\min\{\deg v : v \in F_i\} \geq \max\{\deg v : v \in F_{i+1}\}$$

for all $1 \leq i \leq t-1$.

Remark 7.8 All the components of a greedy forest, except possibly one, have only vertices of degree 1 or 0.

If T is a rooted tree, then we denote its root by $r(T)$. Whenever we consider $T - r(T)$ as rooted forest, we take as root in each connected component the unique neighbor of $r(T)$ contained in the component. The set of all rooted (or edge-rooted) trees with leveled degree sequence D is denoted by $\mathbb{T}_D$. If D is a degree sequence, then $\mathbb{T}_D$ is the set of all trees with degree sequence D. The set of the connected components of $T - r(T)$ is denoted by $C(T)$.

In the remaining part of this chapter, we list a few observations describing properties of sequences, which will then be applied to degree sequences in the following chapters.

We denote by $\mathcal{S}_n$ the set of all permutations of $\{1, \ldots, n\}$.

Let $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ be sequences of non-negative numbers. If A majorizes $(b_{\sigma(1)}, \ldots, b_{\sigma(n)})$ for any $\sigma \in \mathcal{S}_n$, then we write

$$B \blacktriangleleft A. \tag{7.5}$$

Remark 7.9 Let $\sigma \in \mathcal{S}_n$ be such that $b_{\sigma(1)} \geq \cdots \geq b_{\sigma(n)}$. It is easy to see that $(b_{\sigma'(1)}, \ldots, b_{\sigma'(n)}) \blacktriangleleft (b_{\sigma(1)}, \ldots, b_{\sigma(n)})$ for any $\sigma' \in \mathcal{S}_n$. The relation (7.5) is equivalent to the fact that A majorizes $(b_{\sigma(1)}, \ldots, b_{\sigma(n)})$. Furthermore, (7.5) is equivalent to the fact that for any $k \in \{1, \ldots, n\}$ we have

$$(b_{\sigma'(1)}, \ldots, b_{\sigma'(k)}) \blacktriangleleft (a_1, \ldots, a_k)$$

for all $\sigma' \in \mathcal{S}_n$.

Lemma 7.10 ([75]) *Suppose that* $(b_1, \ldots, b_n) \blacktriangleleft (a_1, \ldots, a_n)$ *and* $(b'_1, \ldots, b'_n) \blacktriangleleft (a'_1, \ldots, a'_n)$. *Then we have*

$$b'_1 b_1 + \cdots + b'_n b_n \leq a'_1 a_1 + \cdots + a'_n a_n.$$

The next, stronger looking, lemma is in fact equivalent to Lemma 7.10.

Lemma 7.11 *Suppose that* $(b_1, \ldots, b_n) \blacktriangleleft (a_1, \ldots, a_n)$ *and* $(b'_1, \ldots, b'_n) \blacktriangleleft (a'_1, \ldots, a'_n)$. *Then we have*

$$(b'_1 b_1, \ldots, b'_n b_n) \blacktriangleleft (a'_1 a_1, \ldots, a'_n a_n).$$

Proof. Let σ be the element of $\mathcal{S}_n$ such that $b_{\sigma(1)} b'_{\sigma(1)} \geq \cdots \geq b_{\sigma(n)} b'_{\sigma(n)}$. Using Remark 7.9, we know that for any $k \in \{1, \ldots, n\}$ we have

$$(b_{\sigma(1)}, \ldots, b_{\sigma(k)}) \blacktriangleleft (a_1, \ldots, a_k)$$

and

$$(b'_{\sigma(1)}, \ldots, b'_{\sigma(k)}) \blacktriangleleft (a'_1, \ldots, a'_k).$$

By Lemma 7.10 this implies

$$b_{\sigma(1)} b'_{\sigma(1)} + \cdots + b_{\sigma(k)} b'_{\sigma(k)} \leq a_1 a'_1 + \cdots + a_k a'_k$$

Hence, $(a_1a'_1, \ldots, a_na'_n)$ majorizes $(b_{\sigma(1)}b'_{\sigma(1)}, \ldots, b_{\sigma(n)}b'_{\sigma(n)})$, and the lemma follows from Remark 7.9. $\square$

Let $(k_1, \ldots, k_n)$ be a sequence of nonnegative integers. For any sequence $(a_1, \ldots, a_n)$ we define $(a_1, \ldots, a_n) * (k_1, \ldots, k_n)$ to be the $(k_1 + \cdots + k_n)$-tuple where a_1 is repeated k_1 times and then a_2 is repeated k_2 times, ..., a_n is repeated k_n times. That is

$$(a_1, \ldots, a_n) * (k_1, \ldots, k_n) = \left(b_1, \ldots, b_{\sum_{i=1}^n k_i}\right),$$

where $b_j = a_{j'}$ whenever $\sum_{i=1}^{j'-1} k_i < j \leq \sum_{i=1}^{j'} k_i$. For example $(1, 3, 2) * (2, 3, 4) = (1, 1, 3, 3, 3, 2, 2, 2, 2)$.

Remark 7.12 It is easy to see that if the sequences $(k_1, \ldots, k_n)$ and $(a_1, \ldots, a_n)$ are non-increasing, then for any σ and π in $\mathcal{S}_n$ we have

$$(a_{\sigma(1)}, \ldots, a_{\sigma(n)}) * (k_{\pi(1)}, \ldots, k_{\pi(n)}) \blacktriangleleft (a_1, \ldots, a_n) * (k_1, \ldots, k_n).$$

Lemma 7.13 *Assume that $B = (b_1, \ldots, b_n) \blacktriangleleft (a_1, \ldots, a_n) = A$ and let $C = (c_1, \ldots, c_n)$ be a non-increasing sequence of positive integers. Then for any $\sigma \in \mathcal{S}_n$ we have $B * (c_{\sigma(1)}, \ldots, c_{\sigma(n)}) \blacktriangleleft A * C$.*

Proof. Let $\sigma' \in \mathcal{S}_n$ be such that $b_{\sigma'(1)} \geq \cdots \geq b_{\sigma'(n)}$, and let $B_{\sigma'} = (b_{\sigma'(1)}, \ldots, b_{\sigma'(n)})$. By Remark 7.12, we know that $B * (c_{\sigma(1)}, \ldots, c_{\sigma(n)}) \blacktriangleleft B_{\sigma'} * C$. Since $B_{\sigma'} * C$ is a non-increasing sequence, we can prove the lemma by showing that $A * C$ majorizes $B_{\sigma'} * C$.

The case of $n = 1$ is trivial. Assume that $A * C$ majorizes $B_{\sigma'} * C$ for $n = k$, whenever $B \blacktriangleleft A$. For $n = k + 1$, the relation $B \blacktriangleleft A$ implies that $(b_{\sigma'(1)}, \ldots, b_{\sigma'(k)}) \blacktriangleleft (a_1, \ldots, a_k)$. By the induction hypothesis we deduce that

$$(b_{\sigma'(1)}, \ldots, b_{\sigma'(k)}) * (c_1, \ldots, c_k) \blacktriangleleft (a_1, \ldots, a_k) * (c_1, \ldots, c_k). \tag{7.6}$$

Now we reason by induction with respect to c_{k+1}. For any two sequences $S = (s_1, \ldots, s_l)$ and $S' = (s'_1, \ldots, s'_{l'})$, let $S : S'$ denote the sequence obtained by concatenation, i.e. $S : S' = (s_1, \ldots, s_l, s'_1, \ldots, s'_{l'})$. If $c_{k+1} = 1$, then

$$(b_{\sigma'(1)}, \ldots, b_{\sigma'(k+1)}) * C = ((b_{\sigma'(1)}, \ldots, b_{\sigma'(k)}) * (c_1, \ldots, c_k)) : (b_{\sigma'(k+1)})$$

and $A * C = ((a_1, \ldots, a_k) * (c_1, \ldots, c_k)) : (a_{k+1})$. Using Lemma 7.10 we know that

$$\mathrm{Sum}(B_{\sigma'} * C) = \sum_{i=1}^n b_{\sigma'(i)} c_i \leq \sum_{i=1}^n a_i c_i = \mathrm{Sum}(A * C),$$

where $\mathrm{Sum}(B_{\sigma'} * C)$ and $\mathrm{Sum}(A * C)$ are respectively the sums of the entries in $B_{\sigma'} * C$ and $A * C$. With (7.6) this implies that $A * C$ majorizes

$(b_{\sigma'(1)}, \ldots, b_{\sigma'(k+1)}) * C$. The (second) induction step follows from the relations

$$\begin{aligned} B_{\sigma'} * (c_1, \ldots, c_{k+1}) &= (b_{\sigma'(1)}, \ldots, b_{\sigma'(k+1)}) * (c_1, \ldots, c_{k+1} - 1) : (b_{\sigma'(k+1)}) \\ A * (c_1, \ldots, c_{k+1}) &= (a_1, \ldots, a_{k+1}) * (c_1, \ldots, c_{k+1} - 1) : (a_{k+1}). \end{aligned}$$

$\square$

Chapter 8

Spectral moments of trees with given degree sequence

Let G be a graph with adjacency matrix A, and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. As we already mentioned in Chapter 7, the k-th spectral moment of G is defined as

$$\mathrm{M}_k(G) = \sum_{k=0}^{n} \lambda_i^k. \tag{8.1}$$

A *walk* of length k in a graph G is any sequence $w_1 w_2 \ldots w_{k+1}$ of vertices of G such that $w_i w_{i+1}$ is an edge in G for $i = 1, \ldots, k$. Since $\mathrm{tr}(A^k) = \mathrm{M}_k(G)$, where $\mathrm{tr}(A^k)$ is the trace of the k-th power of A, $\mathrm{M}_k(G)$ is (see [15]) exactly the number of closed walks (walks that start and end at the same vertex) of length k in G. Using the relations

$$\mathrm{EE}(G) = \sum_{i=1}^{n} e^{\lambda_i} = \sum_{i=1}^{n} \sum_{k=0}^{\infty} \frac{\lambda_i^k}{k!} = \sum_{k=0}^{\infty} \frac{\mathrm{M}_k(G)}{k!} \tag{8.2}$$

and

$$\rho(G) = \lim_{k \to \infty} \sqrt[2k]{\mathrm{M}_{2k}(G)}$$

between spectral moments and the *Estrada index* EE, and the spectral radius ρ respectively, we will deduce theorems on EE and ρ from some theorems on M_k. Ernesto Estrada [21] introduced the parameter EE in 2000 and showed how it can be used to study aspects of molecular structures such as the degree of folding of proteins, see also [22, 23]. Applications of EE expanded quickly to the study of complex networks [24] and quantum chemistry [25]. See [32] for a recent survey on the Estrada index.

More generally, for any function $f : \mathbb{R} \to \mathbb{R}$, we define

$$\mathrm{E}_f(G) = \sum_{i=1}^{n} f(\lambda_i). \tag{8.3}$$

Obviously, we obtain the k-th spectral moment for $f(x) = x^k$, the Estrada index for $f(x) = e^x$ and the graph energy (see Part II) for $f(x) = |x|$. More examples will be discussed at a later stage. If we assume that f has a power series expansion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \tag{8.4}$$

around 0 that converges everywhere, then E_f satisfies the relation

$$\mathrm{E}_f(G) = \sum_{i=1}^{n} \sum_{k=0}^{\infty} a_k \lambda_i^k = \sum_{k=0}^{\infty} a_k \mathrm{M}_k(G). \tag{8.5}$$

For any degree sequence D, we prove that $G(D)$ has maximum k-th spectral moment for any $k \geq 0$, and for sufficiently large k, it is unique with this property. Consequently, the greedy tree also maximizes E_f for any f as in (8.4) among all elements of $\mathbb{T}_D$, provided that the coefficients a_k are nonnegative for even k (the odd spectral moments are 0 for all bipartite graphs, thus in particular for trees). Details of the proof are provided in Section 8.1. Furthermore, in Section 8.2 we show that for any degree sequences of trees $B = (b_1, \ldots, b_n) \preccurlyeq (d_1, \ldots, d_n) = D$ we have $\mathrm{M}_k(G(B)) \leq \mathrm{M}_k(G(D))$ for any $k \geq 0$. A number of corollaries can be deduced from these results. In particular a conjecture of Ilić and Stevanović follows as a corollary to our theorems, which reads as follows:

Conjecture 8.1 (Ilić/Stevanović [57]) *For any $k \geq 2$, the Volkmann tree (see Figure 8.1) has maximum spectral moment M_{2k} among trees of n vertices with maximum degree Δ.*

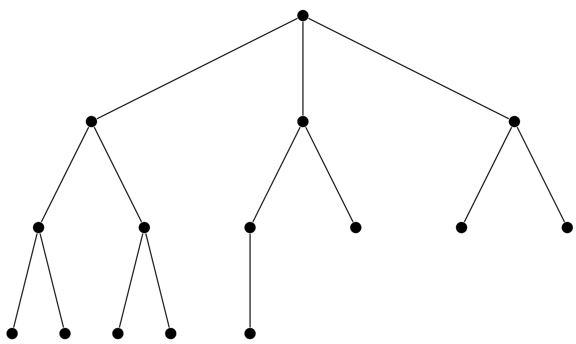


Figure 8.1: The Volkmann tree for $\Delta = 3$, $n = 15$

This, in turn, implies an older conjecture of Gutman, Furtula, Marković and Glišić [35], stating that the Volkmann tree has greatest

Estrada index among all trees with maximum degree Δ, see also [32, p.168]. The Volkmann tree, shown in Figure 8.1 in the case $\Delta = 3$ and $n = 15$, is essentially a complete Δ-ary tree, and a special case of a greedy tree whose degree sequence is $(\Delta, \Delta, \ldots, \Delta, r, 1, 1, \ldots, 1)$ for some r between 1 and Δ. Gutman et al. provide an argument supporting their conjecture, which however is not fully rigorous. The Volkmann tree is well-known to be extremal for other graph invariants, notably for the Wiener index [26].

The conjecture of Ilić and Stevanović was proved by Zhang, Zhou and Li [89] in the case that the maximum degree Δ is large (greater than $n/3$). See [17–20] for further recent extremal results concerning the Estrada index, in particular the Estrada index of trees.

8.1 Trees with given degree sequence

Let T be a tree and v one of its vertices. We denote by $\mathcal{W}_v(k;T)$ the set of all walks of length k in T starting at v, and by $\mathcal{C}_v(k;T)$ the set of all closed walks of length k in T starting and ending at v. We also write

$$\mathcal{W}(k;T) = \bigcup_{v \in V(T)} \mathcal{W}_v(k;T) \tag{8.6}$$

for the set of all walks of length k in T and

$$\mathcal{C}(k;T) = \bigcup_{v \in V(T)} \mathcal{C}_v(k;T) \tag{8.7}$$

for the set of all closed walks of length k. Note that $\mathcal{C}(k;T) = \emptyset$ whenever k is odd.

8.1.1 Vertex rooted trees

Let $W = w_1 \ldots w_k$ be a walk in a rooted tree T. We say that $(i_1, i_2, \ldots, i_k)$ is the *level sequence* of W if w_l is at the i_l^{th} level in T, i.e., at distance $i_l - 1$ from the root, for all $l \leq k$. We denote by $\mathcal{W}(i_1, \ldots, i_k;T)$ the set of walks with level sequence $(i_1, \ldots, i_k)$ in T. For any vertex v of T we define

$$\mathcal{W}_v(i_1, \ldots, i_k;T) = \{w_1 \ldots w_k \in \mathcal{W}(i_1, \ldots, i_k;T) : w_1 = v\}.$$

$\mathcal{C}(i_1, \ldots, i_k;T)$ and $\mathcal{C}_v(i_1, \ldots, i_k;T)$ are the subsets of $\mathcal{W}(i_1, \ldots, i_k;T)$ and $\mathcal{W}_v(i_1, \ldots, i_k;T)$, respectively, which contain the closed walks. Moreover, we denote the cardinalities of $\mathcal{W}(k;T)$ and $\mathcal{C}(k;T)$ by $W(k;T)$ and $C(k;T)$ respectively, the cardinality of $\mathcal{W}_v(i_1, \ldots, i_k;T)$ by $W_v(i_1, \ldots, i_k;T)$, etc. This convention will be kept even if not mentioned explicitly.

Lemma 8.2 *Let $T \in \mathbb{T}_D$ for some leveled degree sequence D of a vertex-rooted forest, and let $G = G(D)$ be the associated level greedy forest. Let $v_1^i, \ldots, v_{d_i}^i$ be the vertices of T at the i^{th} level. Then for any level sequence of walks $(i_1, \ldots, i_l)$, the following relations hold for all i:*

$$\begin{aligned}&(W_{v_1^i}(i_1,\ldots,i_l;T),\ldots,W_{v_{d_i}^i}(i_1,\ldots,i_l;T))\\ &\qquad\blacktriangleleft (W_{g_1^i}(i_1,\ldots,i_l;G),\ldots,W_{g_{d_i}^i}(i_1,\ldots,i_l;G))\end{aligned} \tag{8.8}$$

and

$$W_{g_1^i}(i_1,\ldots,i_l;G) \geq W_{g_2^i}(i_1,\ldots,i_l;G) \geq \cdots \geq W_{g_{d_i}^i}(i_1,\ldots,i_l;G). \tag{8.9}$$

Proof. The situation where $i \neq i_1$ is not interesting, since we get

$$W_{v_j^i}(i_1,\ldots,i_l;T) = W_{g_j^i}(i_1,\ldots,i_l;G) = 0$$

for any j. So we assume that $i = i_1$ and proceed by induction with respect to l. The initial case $l = 1$ is trivial, since we know that

$$W_{v_j^{i_1}}(i_1;T) = W_{g_j^{i_1}}(i_1;G) = 1$$

for all i_1 and j. Assume that the relations (8.8) and (8.9) hold whenever $l \leq k$ for some integer $k \geq 1$. Now consider a longer level sequence $(i_1, \ldots, i_l)$ where $l = k+1$. Note that by the induction hypothesis we have

$$\begin{aligned}&(W_{v_1^m}(i_2,\ldots,i_l;T),\ldots,W_{v_{d_i}^m}(i_2,\ldots,i_l;T))\\ &\qquad\blacktriangleleft (W_{g_1^m}(i_2,\ldots,i_l;G),\ldots,W_{g_{d_m}^m}(i_2,\ldots,i_l;G))\end{aligned} \tag{8.10}$$

and

$$W_{g_1^m}(i_2,\ldots,i_l;G) \geq W_{g_2^m}(i_2,\ldots,i_l;G) \geq \cdots \geq W_{g_{d_m}^m}(i_2,\ldots,i_l;G) \tag{8.11}$$

for any level m. There are two cases: $i_2 = i_1 - 1$ or $i_2 = i_1 + 1$ (in all other cases, the number of walks is 0).

Case 1: Assume that $i_2 = i_1 + 1 = i + 1$. For $1 \leq j \leq d_i$, we use a_j as an abbreviation for the number of children of v_j^i and b_j for the number of children of g_j^i. Clearly, $a_j = \deg v_j^i - 1$ and $b_j = \deg g_j^i - 1$ if $i \neq 1$, and $a_j = \deg v_j^i$, $b_j = \deg g_j^i$ if $i = 1$. In view of the construction of greedy trees, we have

$$b_1 \geq \cdots \geq b_{d_i}, \tag{8.12}$$

and since $(a_1, \ldots, a_{d_i})$ is a permutation of $(b_1, \ldots, b_{d_i})$, it is clear that

$$(a_1, \ldots, a_{d_i}) \blacktriangleleft (b_1, \ldots, b_{d_i}). \tag{8.13}$$

We also write r_j and s_j for the sums

$$r_j = \sum_{t=1}^{j} a_t \quad \text{and} \quad s_j = \sum_{t=1}^{j} b_t,$$

and $r_0 = s_0 = 0$. Now note that

$$W_{v_j^i}(i_1, \ldots, i_l; T) = \sum_{v_h^{i+1} \sim v_j^i} W_{v_h^{i+1}}(i_2, \ldots, i_l; T) = \sum_{h=r_{j-1}+1}^{r_j} W_{v_h^{i+1}}(i_2, \ldots, i_l; T),$$

since every walk with level sequence $(i_1, \ldots, i_l)$ starting at v_j^i has to go to one of the children v_h^{i+1} ($r_{j-1} + 1 \le h \le r_j$) first. Likewise,

$$W_{g_j^i}(i_1, \ldots, i_l; G) = \sum_{g_h^{i+1} \sim g_j^i} W_{g_h^{i+1}}(i_2, \ldots, i_l; G) = \sum_{h=s_{j-1}+1}^{s_j} W_{g_h^{i+1}}(i_2, \ldots, i_l; G).$$

With (8.11) and (8.12) this implies (8.9) for $l = k + 1$. Furthermore, the majorization (8.8) for $l = k + 1$ follows from (8.10) and (8.13).

Case 2: Assume that $i_2 = i_1 - 1 = i - 1$. This time, we write a_j for the number of children of v_j^{i-1} (which is either $\deg v_j^{i-1}$ or $\deg v_j^{i-1} - 1$) and b_j for the number of children of g_j^{i-1}. The relation (8.13) is still valid. Now we have

$$\begin{aligned}&(W_{v_1^i}(i_1, \ldots, i_l; T), \ldots, W_{v_{d_i}^i}(i_1, \ldots, i_l; T)) \\ &\qquad = (W_{v_1^{i-1}}(i_2, \ldots, i_l; T), \ldots, W_{v_{d_{i-1}}^{i-1}}(i_2, \ldots, i_l; T)) * (a_1, \ldots, a_{d_{i-1}}),\end{aligned}$$

since if v_h^i is one of the a_j children of v_j^{i-1}, a walk with level sequence $(i_1, \ldots, i_l)$ starting at v_h^i has to start with a step to v_j^{i-1}, which means that

$$W_{v_h^i}(i_1, \ldots, i_l; T) = W_{v_j^{i-1}}(i_2, \ldots, i_l; T).$$

Likewise,

$$\begin{aligned}&(W_{g_1^i}(i_1, \ldots, i_l; G), \ldots, W_{g_{d_i}^i}(i_1, \ldots, i_l; G)) \\ &= (W_{g_1^{i-1}}(i_2, \ldots, i_l; G), \ldots, W_{g_{d_{i-1}}^{i-1}}(i_2, \ldots, i_l; G)) * (b_1, \ldots, b_{d_{i-1}}). \qquad (8.14)\end{aligned}$$

This shows that (8.9) for $l = k + 1$ is a direct consequence of (8.11), and (8.8) for $l = k + 1$ follows from (8.13) and (8.10) by means of Lemma 7.13. $\square$

Next we study closed walks: it turns out that a completely analogous statement holds.

Lemma 8.3 *Let $T \in \mathbb{T}_D$ for some leveled degree sequence D of a vertex-rooted forest, and let $G = G(D)$ be the associated greedy forest. Let $v_1^i, \dots, v_{d_i}^i$ be the vertices of T at the i^{th} level. Then for any level sequence of walks $(i_1, \dots, i_l)$, the following relations hold for all i:*

$$\begin{aligned}&(C_{v_1^i}(i_1, \dots, i_l; T), \dots, C_{v_{d_i}^i}(i_1, \dots, i_l; T)) \\ &\qquad \blacktriangleleft (C_{g_1^i}(i_1, \dots, i_l; G), \dots, C_{g_{d_i}^i}(i_1, \dots, i_l; G))\end{aligned} \tag{8.15}$$

and

$$C_{g_1^i}(i_1, \dots, i_l; G) \geq \dots \geq C_{g_{d_i}^i}(i_1, \dots, i_l; G). \tag{8.16}$$

Proof. As in the proof of Lemma 8.2, we only need to prove the lemma for $i = i_1$. The case when l is even is trivial: in this case,

$$\mathcal{C}_{v_j^i}(i_1, \dots, i_l; T) = \mathcal{C}_{g_j^i}(i_1, \dots, i_l; G) = \emptyset$$

for all j, since there are no closed walks of odd length in a forest.

For the case of odd l, say $l = 2l' - 1$, the proof is similar to that of Lemma 8.2: We reason by induction with respect to l'. The case $l' = 1$ is again trivial. Assume that the lemma holds for all $l' \leq k$ for some $k \geq 1$. Now consider a level sequence $(i_1, \dots, i_{2k+1})$. The induction hypothesis implies that

$$\begin{aligned}&(C_{v_1^i}(i_m, \dots, i_{2l'+1}; T), \dots, C_{v_{d_i}^i}(i_m, \dots, i_{2l'+1}; T)) \\ &\qquad \blacktriangleleft (C_{g_1^i}(i_m, \dots, i_{2l'+1}; G), \dots, C_{g_{d_i}^i}(i_m, \dots, i_{2l'+1}; G))\end{aligned} \tag{8.17}$$

and

$$C_{g_1^i}(i_m, \dots, i_{2l'+1}; G) \geq \dots \geq C_{g_{d_i}^i}(i_m, \dots, i_{2l'+1}; G) \tag{8.18}$$

for any $1 < m \leq 2l' + 1$. We also must have $i_1 = i_{2k+1} = i$ and $i_2 = i \pm 1$ as well as $i_{2k} = i \pm 1$, the other possibilities are trivial.

Case 1: If $i_2 = i_{2k} = i - 1$, then, writing a_j for the number of children of v_j^{i-1} and b_j for the number of children of g_j^{i-1}, we have

$$\begin{aligned}&(C_{v_1^i}(i_1, \dots, i_{2k+1}; T), \dots, C_{v_{d_i}^i}(i_1, \dots, i_{2k+1}; T)) \\ &= (C_{v_1^{i-1}}(i_2, \dots, i_{2k}; T), \dots, C_{v_{d_{i-1}}^{i-1}}(i_2, \dots, i_{2k}; T)) * (a_1, \dots, a_{d_{i-1}})\end{aligned}$$

and

$$\begin{aligned}&(C_{g_1^i}(i_1, \dots, i_{2k+1}; G), \dots, C_{g_{d_i}^i}(i_1, \dots, i_{2k+1}; G)) \\ &= (C_{g_1^{i-1}}(i_2, \dots, i_{2k}; G), \dots, C_{g_{d_{i-1}}^{i-1}}(i_2, \dots, i_{2k}; G)) * (b_1, \dots, b_{d_{i-1}})\end{aligned}$$

for the same reason as in Case 2 of Lemma 8.2. Hence (8.16) follows from (8.12) and (8.18), and (8.15) can be obtained from (8.13) and (8.17) using Lemma 7.13.

Case 2: Assume that $i_2 = i_1 + 1 = i + 1$. Let h be the smallest integer such that $h > 1$ and $i_h = i_1 = i$. If h does not exist, then there is no closed walk with level sequence $(i_1, \ldots, i_{2k+1})$, so we can ignore this case. By the definition of h and the assumption that $i_2 = i + 1$, we know that $i = \min\{i_1, \ldots, i_h\}$. Clearly, any walk with level sequence $(i_1, \ldots, i_h)$ is closed. Hence for all j, any element of $\mathcal{C}_{g_j^i}(i_1, \ldots, i_{2k+1}; G)$ can be decomposed (uniquely) into a first part that is an element of $\mathcal{C}_{g_j^i}(i_1, \ldots, i_h; G)$ and a second part that is an element of $\mathcal{C}_{g_j^i}(i_h, \ldots, i_{2k+1}; G)$. Similarly, an element of $\mathcal{C}_{v_j^i}(i_1, \ldots, i_{2k+1}; T)$ splits (uniquely) into two parts: a first part in $\mathcal{C}_{v_j^i}(i_1, \ldots, i_h; T)$ and a second part in $\mathcal{C}_{v_j^i}(i_h, \ldots, i_{2k+1}; T)$. This implies that

$$\begin{aligned} C_{g_j^i}(i_1, \ldots, i_{2k+1}; G) &= C_{g_j^i}(i_1, \ldots, i_h; G) C_{g_j^i}(i_h, \ldots, i_{2k+1}; G) \\ &= W_{g_j^i}(i_1, \ldots, i_h; G) C_{g_j^i}(i_h, \ldots, i_{2k+1}; G) \end{aligned}$$

and

$$\begin{aligned} C_{v_j^i}(i_1, \ldots, i_{2k+1}; T) &= C_{v_j^i}(i_1, \ldots, i_h; T) C_{v_j^i}(i_h, \ldots, i_{2k+1}; T) \\ &= W_{v_j^i}(i_1, \ldots, i_h; T) C_{v_j^i}(i_h, \ldots, i_{2k+1}; T). \end{aligned}$$

Therefore, we can use Lemma 8.2, the induction hypothesis and Lemma 7.11 to deduce (8.15) and (8.16); the argument remains valid even if $h = 2k + 1$, since then the second factor in the formulas above is simply 1.

Case 3: Assume that $i_{2k} = i_{2k+1} + 1 = i + 1$. Then the sequence $(i_{2k+1}, \ldots, i_1)$ satisfies the condition of Case 2. Hence, for this case, (8.15) and (8.16) follow from the fact that for any j we have

$$C_{g_j^i}(i_1, \ldots, i_{2k+1}; G) = C_{g_j^i}(i_{2k+1}, \ldots, i_1; G)$$

and

$$C_{v_j^i}(i_1, \ldots, i_{2k+1}; T) = C_{v_j^i}(i_{2k+1}, \ldots, i_1; T).$$

This completes the proof, since there are no closed walks in any other cases. □

The following theorem is a direct consequence of the two Lemmas 8.2 and 8.3 and the relations (8.6) and (8.7).

Theorem 8.4 *Let D be a leveled degree sequence of a vertex-rooted forest and $G(D)$ the associated level greedy forest. Then for any non-negative integer k and all $T \in \mathbb{T}_D$, we have*

$$W(k; T) \leq W(k; G(D))$$

and

$$\mathrm{M}_k(T) = C(k; T) \leq C(k; G(D)) = \mathrm{M}_k(G(D)).$$

It turns out that one has strict inequality for sufficiently large even k, which is shown in the following lemma:

Lemma 8.5 *Let D be a leveled degree sequence of a vertex-rooted forest and $G = G(D)$ the associated level greedy forest. If $T \in \mathbb{T}_D$ is not isomorphic (as a rooted forest) to G, then there exists an integer k_0 such that*

$$\mathrm{M}_k(T) = C(k;T) < C(k;G(D)) = \mathrm{M}_k(G(D))$$

for all even $k \geq k_0$.

Proof. It suffices to find one specific level sequence for which we have strict inequality. We take h_2 to be the smallest positive integer such that T, restricted to the first h_2 levels, is not isomorphic to a level greedy rooted forest. Then let h_1 be the largest positive integer such that the restriction of T to levels $h_1, h_1 + 1, \ldots, h_2$ (which we denote by P, see Figure 8.2 for an example) is still not isomorphic to a greedy rooted forest.

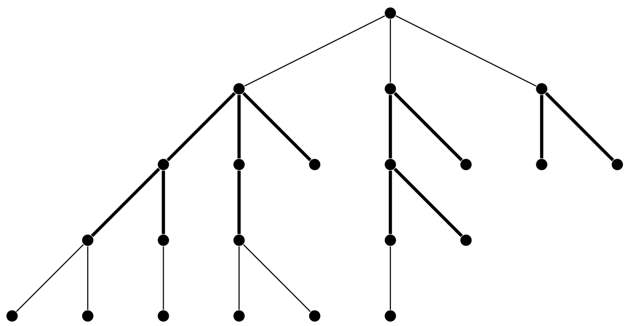


Figure 8.2: Example of the forest P described in the proof of Lemma 8.5 for a given T: $h_1 = 2, h_2 = 4$ and the bold subforest is P

From now on, we only work with the restricted forest P. Let r be the number of its roots and $P_1, P_2, \ldots, P_r$ the components of P. Each of them is a level greedy tree: if not, we could remove the roots to obtain a rooted forest that is not level greedy, contradicting the maximality of h_1. However, by assumption, their union is not a level greedy forest.

Now let $p_1, p_2, \ldots, p_r$ be the number of descendants of the r roots at level h_2 (p_j descendants in component P_j). The analogous numbers for the greedy tree are $q_1, q_2, \ldots, q_r$, and we call the corresponding components of the restriction of G to the same levels $Q_1, Q_2, \ldots, Q_r$.

We assume, without loss of generality, that $p_1 \geq p_2 \geq \cdots \geq p_r$ and $q_1 \geq q_2 \geq \cdots \geq q_r$. From the construction of level greedy forests, we know that

$$(p_1, p_2, \ldots, p_r) \blacktriangleleft (q_1, q_2, \ldots, q_r). \tag{8.19}$$

In fact, this is a special case of Lemma 8.2, since $p_1, \ldots, p_r$ and $q_1, \ldots, q_r$ also count walks with level sequence $(h_1, h_1 + 1, \ldots, h_2)$. The number of closed walks with level sequence

$$(h_2, h_2 - 1, \ldots, h_1 + 1, h_1, h_1 + 1, \ldots, h_2 - 1, h_2, h_2 - 1, \\ \ldots, h_1 + 1, h_1, h_1 + 1, \ldots, h_2) \tag{8.20}$$

in T and G are

$$p_1^2 + p_2^2 + \cdots + p_r^2 \quad \text{and} \quad q_1^2 + q_2^2 + \cdots + q_r^2$$

respectively: such walks start at level h_2, move up to the root, return to level h_2, then back to the root, and back to the starting point. They are thus completely determined by the two vertices at level h_2 (not necessarily distinct), which have to have the same root.

We suppose first that $p = (p_1, p_2, \ldots, p_r) \neq (q_1, q_2, \ldots, q_r) = q$. Let i be the first index and j the last index where the two differ. Since q majorizes p and the two have the same sums, we must have $q_i > p_i$ and $q_j < p_j$. Let $\epsilon = \min\{q_i - p_i, p_j - q_j\}$, and replace p_i by $p_i + \epsilon$ and p_j by $p_j - \epsilon$. Then the sum of squares increases by

$$(p_i + \epsilon)^2 - p_i^2 + (p_j - \epsilon)^2 - p_j^2 = 2\epsilon(p_i - p_j + \epsilon) > 0.$$

Repeating this process, we can transform p into q, which shows that

$$p_1^2 + p_2^2 + \cdots + p_r^2 < q_1^2 + q_2^2 + \cdots + q_r^2, \tag{8.21}$$

and we are done in that we have found a level sequence such that G has strictly more closed walks than T. This is for closed walks of length $k_0 := 4(h_2 - h_1)$. For any even $k > k_0$ we consider level sequences of the form

$$(h_2, h_2 - 1, \ldots, h_1 + 1, h_1, \boldsymbol{h_1 + 1, h_1, h_1 + 1, h_1, \ldots,} \\ \boldsymbol{h_1}, h_1 + 1, \ldots, h_2, h_2 - 1, \ldots, h_1, h_1 + 1, \ldots, h_2),$$

similar to (8.20) except that we introduced a $k - k_0$ entries subsequence (in bold). The number of closed walks with such a level sequence is

$$N := p_1^2 s_1^{(k-k_0)/2} + p_2^2 s_2^{(k-k_0)/2} + \cdots + p_r^2 s_r^{(k-k_0)/2}$$

in T and

$$M := q_1^2 t_1^{(k-k_0)/2} + q_2^2 t_2^{(k-k_0)/2} + \cdots + q_r^2 t_r^{(k-k_0)/2}$$

in G, where $s_i = \deg(\mathrm{r}(P_i))$ and $t_i = \deg(\mathrm{r}(Q_i))$ for any $1 \leq i \leq r$. Using Lemma 7.11, we deduce from (8.19) and

$$(s_1, s_2, \cdots, s_r) \blacktriangleleft (t_1, t_2, \cdots, t_r) \tag{8.22}$$

that

$$(p_1^2, p_2^2, \cdots, p_r^2) \blacktriangleleft (q_1^2, q_2^2, \cdots, q_r^2), \tag{8.23}$$

and

$$(s_1^{(k-k_0)/2}, s_2^{(k-k_0)/2}, \cdots, s_r^{(k-k_0)/2}) \blacktriangleleft (t_1^{(k-k_0)/2}, t_2^{(k-k_0)/2}, \cdots, t_r^{(k-k_0)/2}). \tag{8.24}$$

(8.21) and (8.23) imply that

$$(p_1^2, p_2^2, \cdots, p_r^2) \blacktriangleleft (q_1^2, q_2^2, \cdots, q_r^2 - 1). \tag{8.25}$$

With use of Lemma 7.10, (8.24) and (8.25) imply $N < M$. This completes the proof in the case that p and q are not identical.

Let us now assume that $p = (p_1, p_2, \ldots, p_r) = (q_1, q_2, \ldots, q_r) = q$, and let l be the last index such that $p_l = q_l \neq 0$. By our choice of h_2, the restrictions of T and G to levels $h_1, h_1 + 1, \ldots, h_2 - 1$ are isomorphic: they are both level greedy forests consisting of r components. If one component is larger than another, then the number of vertices at level $h_2 - 1$ is greater as well, and if two components have the same number of vertices at level $h_2 - 1$, then they are isomorphic by the construction of greedy trees.

Let m be the number of vertices at level $h_2 - 1$ in the largest component. Then q_1 is the sum of the highest m degrees at level $h_2 - 1$. The only way how p_1 can be equal to q_1 is thus that P_1 and Q_1 have the same number of vertices at level $h_2 - 1$, so they have to be isomorphic (both are known to be level greedy as well!). Likewise, P_2 and Q_2 have to be isomorphic, etc. The only possible exception are P_l and Q_l, the last components with vertices at level h_2: here, some vertices in Q_l at level $h_2 - 1$ might be leaves, so P_l could be smaller than Q_l.

Now let $p_1', p_2', \ldots, p_r', q_1', q_2', \ldots, q_r'$ be the number of vertices at level $h_2 - 1$ in $P_1, P_2, \ldots, P_r, Q_1, Q_2, \ldots, Q_r$ respectively. The number of closed walks with level sequence

$$(h_2 - 1, \ldots, h_1 + 1, h_1, \boldsymbol{h_1 + 1, h_1, h_1 + 1, h_1, \ldots,}$$
$$\boldsymbol{h_1}, h_1 + 1, \ldots, h_2, h_2 - 1, \ldots, h_1, h_1 + 1, \ldots, h_2 - 1),$$

where the bold section has $a := k - (4(h_2 - h_1) - 2) \geq 0$ entries for some even k, in T and G are

$$p_1 p_1' s_1^{a/2} + p_2 p_2' s_2^{a/2} + \cdots + p_r p_r' s_r^{a/2}$$

and

$$q_1 q_1' t_1^{a/2} + q_2 q_2' t_2^{a/2} + \cdots + q_r q_r' t_r^{a/2}$$

respectively, by the same reasoning as before. We know that $p_i p_i' = q_i q_i'$ and $s_i^{a/2} = t_i^{a/2}$ for $i < l$, thus for (8.22) to hold we must have

$s_l \leq t_l$; furthermore we also know that $p_i = q_i = 0$ for $i > l$. Hence we have

$$\begin{aligned}&(q_1 q_1' t_1^{a/2} + q_2 q_2' t_2^{a/2} + \cdots + q_r q_r' t_r^{a/2}) - (p_1 p_1' s_1^{a/2} + p_2 p_2' s_2^{a/2} + \cdots + p_r p_r' s_r^{a/2})\\ &= q_l q_l' t_l^{a/2} - p_l p_l' s_l^{a/2} = q_l (q_l' t_l^{a/2} - p_l' s_l^{a/2}).\end{aligned}$$

If $q_l' = p_l'$, then the components P_l and Q_l up to level $h_2 - 1$ have to be isomorphic, and since both are level greedy up to level h_2 as well, they must be isomorphic. But then T and G, restricted to levels $h_1, h_1 + 1, \ldots, h_2$, are isomorphic, contradicting our choice of h_1 and h_2. Thus $q_l' > p_l'$, which means that we have again found a suitable level sequence for all even $k \geq k_0 := 4(h_2 - h_1) - 2$. □

8.1.2 Edge rooted trees

As we will see at the end of this subsection, Theorem 8.4 still holds if we consider edge-rooted trees instead of vertex-rooted trees.

For any set $\mathcal{A}$ of walks in a graph and any vertex v and edge e of the same graph, we denote by $\mathcal{A}^e$ and by $\mathcal{A}^v$ the subsets of $\mathcal{A}$ that only contain walks passing through e and v, respectively. Instead of $(\mathcal{A}^e)^{e'}$ we simply write $\mathcal{A}^{e,e'}$. Similarly, $(\mathcal{A}^e)^v = (\mathcal{A}^v)^e = \mathcal{A}^{v,e} = \mathcal{A}^{e,v}$. For any two adjacent vertices u and v in a graph G, we define $\mathcal{C}_{u,v}(k;G)$ to be the set and $C_{u,v}(k;G)$ the number of all closed walks of length k starting from the edge uv in direction from u to v.

Different combinations of these notations are possible. For example, for some edge uv in a graph G and another edge e, $\mathcal{C}^e_{u,v}(k;G)$ stands for the set of closed walks of length k in G starting at u, using the edge uv at the first step and passing through e at a later stage.

Lemma 8.6 *Let u and v be two adjacent vertices in a graph G, and let e be an edge in G. Then for all nonnegative integers k we have*

$$C_{u,v}(k;G) = C_{v,u}(k;G) \quad \textit{and} \quad C^e_{u,v}(k;G) = C^e_{v,u}(k;G).$$

Proof. Both $C_{u,v}(k;G)$ and $C_{v,u}(k;G)$ are equal to the number of walks of length $k-1$ starting from u and ending at v (which is clearly the same as the number of walks of length $k-1$ starting from v and ending at u).

If $e \neq uv$, then both $C^e_{u,v}(k;G)$ and $C^e_{v,u}(k;G)$ are equal to the number of walks of length $k-1$ starting from u, passing through e and ending at v. If $e = uv$, then clearly $C^e_{u,v}(k;G) = C_{u,v}(k;G)$, and we are done. □

We extend the notation $\mathcal{C}_v(k;T)$ and denote by $\mathcal{C}_e(k;T)$ the set of walks of length k in T which start with the edge e (in either direction).

As usual, $C_v(k;T)$ and $C_e(k;T)$ denote their cardinalities. If T is an edge-rooted tree such that u and v are the ends of $\mathrm{r}(T)$, we know by Lemma 8.6 that

$$C_{\mathrm{r}(T)}(k;T) = C_{u,v}(k;T) + C_{v,u}(k;T) = 2C_{u,v}(k;T) = 2C_{v,u}(k;T).$$

Lemma 8.7 *Let D be a leveled degree sequence of an edge-rooted tree and $G = G(D)$ the associated edge-rooted greedy tree. For any element $T \in \mathbb{T}_D$ we have*

$$C_{\mathrm{r}(T)}(k;T) \le C_{\mathrm{r}(G)}(k;G)$$

for any nonnegative integer k.

Proof. Let G_1 and G_2 be the components of $G - \mathrm{r}(G)$, and let T_1 and T_2 be the components of $T - \mathrm{r}(T)$. Since for odd k we trivially have $C_{\mathrm{r}(T)}(k;T) = C_{\mathrm{r}(G)}(k;G) = 0$, we are only interested in even $k = 2l$. Let us reason by induction on l. The cases where $l = 1, 2$ are easy to check, since the closed walks of length at most 4 starting with the root edge cannot reach beyond the first two levels, but these parts of T and G are isomorphic edge-rooted trees. Assume that the lemma holds whenever $l \le m$ for some integer $m \ge 2$. Now consider the case where $l = m+1$. The level sequences of the elements in $\mathcal{C}_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}_{\mathrm{r}(G)}(k;G)$ are of the form $(1, 1, i_1, i_2, \ldots, i_{k-1})$, and i_{k-1} also has to be 1.

We first consider walks that do not return immediately to the starting point after the first step. For any j with $2 \le j \le k-1$, let $\mathcal{C}^j_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}^j_{\mathrm{r}(G)}(k;G)$ be respectively the subsets of $\mathcal{C}_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}_{\mathrm{r}(G)}(k;G)$ whose elements are the walks with level sequences $(1, 1, i_1, i_2, \ldots, i_{k-1})$, where $i_j = 1$ and $1 \notin \{i_1, i_2, \ldots, i_{j-1}\}$. Their cardinalities are denoted by $C^j_{\mathrm{r}(T)}(k;T)$ and $C^j_{\mathrm{r}(G)}(k;G)$ respectively. These walks start with the edge root, then go on to higher levels, return to level 1 for the first time after j steps, and then continue with $k-j-1$ more steps until they return to the starting point. We can uniquely split each of these walks into the j steps from the first step to level 2 to the first return to level 1 and the rest. Set

$$\mathbb{S}_j = \{(1, i_1, i_2, \ldots, i_{j-1}, 1) : 1 \notin \{i_1, i_2, \ldots, i_{j-1}\}\}.$$

From Lemma 8.3, Lemma 8.6 and the induction hypothesis, we now obtain

$$\begin{aligned}
C^j_{\mathrm{r}(T)}(k;T) &= C_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(k-j;T) \sum_{S \in \mathbb{S}_j} C_{\mathrm{r}(T_2)}(S;T_2) \\
&\quad + C_{\mathrm{r}(T_2),\mathrm{r}(T_1)}(k-j;T) \sum_{S \in \mathbb{S}_j} C_{\mathrm{r}(T_1)}(S;T_1) \\
&= \frac{1}{2} C_{\mathrm{r}(T)}(k-j;T) \Big(\sum_{S \in \mathbb{S}_j} C_{\mathrm{r}(T_2)}(S;T_2) + \sum_{S \in \mathbb{S}_j} C_{\mathrm{r}(T_1)}(S;T_1) \Big)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{2} C_{\mathrm{r}(G)}(k-j;G)\Big(\sum_{S\in\mathbb{S}_j} C_{\mathrm{r}(G_2)}(S;G_2) + \sum_{S\in\mathbb{S}_j} C_{\mathrm{r}(G_1)}(S;G_1)\Big)\\
&= C^j_{\mathrm{r}(G)}(k;G)
\end{aligned}
$$

for any $j \geq 2$. This covers all the cases where $i_1 \neq 1$. Next, consider the subsets $\mathcal{C}^*_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}^*_{\mathrm{r}(G)}(k;G)$ of $\mathcal{C}_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}_{\mathrm{r}(G)}(k;G)$, respectively; their elements are closed walks with level sequence $(1,1,i_1,i_2,\ldots,\ i_{k-1})$, where $i_1 = 1$ and for any $h \in \{1,2,\ldots,k-2\}$ we always have $(1,1) \neq (i_h,i_{h+1})$. In words, these walks move forwards and backwards along the edge root for the first two steps, then never use the edge root again, thus they stay in one of the two branches. From Lemma 8.3, we now get

$$
\begin{aligned}
C^*_{\mathrm{r}(T)}(k;T) &= C_{\mathrm{r}(T_1)}(k-2;T_1) + C_{\mathrm{r}(T_2)}(k-2;T_2)\\
&\leq C_{\mathrm{r}(G_1)}(k-2;G_1) + C_{\mathrm{r}(G_2)}(k-2;G_2)\\
&= C^*_{\mathrm{r}(G)}(k;G). \qquad (8.26)
\end{aligned}
$$

We are left with walks that use the edge root, return immediately, and use the edge root again at some stage. The set of these walks is divided further, depending on the first time that the edge root is used again. For any $j \geq 1$, we consider the subsets $\mathcal{C}'^j_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}'^j_{\mathrm{r}(G)}(k;G)$ of $\mathcal{C}_{\mathrm{r}(T)}(k;T)$ and $\mathcal{C}_{\mathrm{r}(G)}(k;G)$ whose elements are the closed walks with level sequence $(1,1,\ i_1,i_2,\ldots,i_{k-1})$, where $i_1 = i_j = i_{j+1} = 1$ and $(1,1) \neq (i_h,i_{h+1})$ for any $h \in \{1,\ldots,j-1\}$. Such a walk can be split uniquely into a walk of length $j+1$ in $\mathcal{C}^*_{\mathrm{r}(T)}(j+1;T)$ ($\mathcal{C}^*_{\mathrm{r}(G)}(j+1;G)$, respectively) and a closed walk of length $k-j-1$ starting with the edge root. From (8.26) and Lemma 8.6, we obtain

$$
\begin{aligned}
C'^j_{\mathrm{r}(T)}(k;T) &= C^*_{\mathrm{r}(T)}(j+1;T)\cdot \frac{1}{2} C_{\mathrm{r}(T)}(k-j-1;T)\\
&\leq C^*_{\mathrm{r}(G)}(j+1;G)\cdot \frac{1}{2} C_{\mathrm{r}(G)}(k-j-1;G)\\
&= C'^j_{\mathrm{r}(G)}(k;G)
\end{aligned}
$$

for any $j \geq 1$. We see that the greedy tree G has more or at least equally many walks of each type as T, which completes the proof. □

Lemma 8.8 *Let D be a leveled degree sequence of an edge-rooted tree and $G = G(D)$ the associated edge-rooted greedy tree. For any element $T \in \mathbb{T}_D$ and for any nonnegative integer k we have*

$$C^{\mathrm{r}(T)}(k;T) \leq C^{\mathrm{r}(G)}(k;G).$$

Proof. Any element, say W, in $\mathcal{C}^{\mathrm{r}(T)}(k;T)$ or $\mathcal{C}^{\mathrm{r}(G)}(k;G)$ has a unique decomposition as $W = W_1W_2W_3$ for some W_1, W_2 and W_3 satisfying the following conditions:

i) W_2 is a closed walk starting from the edge root, chosen to have maximal length.

ii) W_1 and W_3 do not use the edge root but can possibly have length zero. By merging the end of W_1 with the beginning of W_3, we obtain a closed walk W'.

Under the conditions *i*) and *ii*), W_3 visits an end of the edge root only once (at its starting point), otherwise we could extend W_2. This means that W can be uniquely recovered from W' and W_2 by inserting W_2 into W' at the last appearance of an end vertex of the edge root. So the number of possible walks W is the number of possible walks W' times the number of possible walks W_2.

Let T_1 and T_2 be the components of $T - \mathrm{r}(T)$, and G_1 and G_2 those of $G - \mathrm{r}(G)$. By Lemma 8.3, we know that

$$C^{\mathrm{r}(T_1)}(l;T_1) + C^{\mathrm{r}(T_2)}(l;T_2) \le C^{\mathrm{r}(G_1)}(l;G_1) + C^{\mathrm{r}(G_2)}(l;G_2)$$

for any nonnegative integer l. Hence, using Lemma 8.7 we have

$$\begin{aligned}
&C^{\mathrm{r}(T)}(k;T)\\
&= \sum_{k_1+k_2=k} \big(C^{\mathrm{r}(T_1)}(k_1;T_1)C_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(k_2;T) + C^{\mathrm{r}(T_2)}(k_1;T_2)C_{\mathrm{r}(T_2)\,\mathrm{r}(T_1)}(k_2;T)\big)\\
&= \sum_{k_1+k_2=k} \big(C^{\mathrm{r}(T_1)}(k_1;T_1) + C^{\mathrm{r}(T_2)}(k_1;T_2)\big)\cdot\frac{1}{2}C_{\mathrm{r}(T)}(k_2;T)\\
&\le \sum_{k_1+k_2=k} \big(C^{\mathrm{r}(G_1)}(k_1;G_1) + C^{\mathrm{r}(G_2)}(k_1;G_2)\big)\cdot\frac{1}{2}C_{\mathrm{r}(G)}(k_2;G)\\
&= C^{\mathrm{r}(G)}(k;G).
\end{aligned}$$

□

The next theorem combines Theorem 8.4 and Lemma 8.8.

Theorem 8.9 *Let D be a leveled degree sequence of an edge-rooted tree. For any nonnegative integer k and all $T \in \mathbb{T}_D$, we have*

$$\mathrm{M}_k(T) = C(k;T) \le C(k;G(D)) = \mathrm{M}_k(G(D)).$$

For sufficiently large even k, the inequality is strict unless T and $G(D)$ are isomorphic.

Proof. Use Theorem 8.4 to compare the number of closed walks of length k not using the edge root, and Lemma 8.8 for those which pass through the edge root. The fact that the inequality in Theorem 8.4 is strict for sufficiently large k by Lemma 8.5 implies that this is also the case here. □

8.1.3 Main result

The main result of this section is the fact that if we fix a degree sequence D, then among all trees with degree sequence D, the greedy tree $G(D)$ has the maximum number of closed walks of any given length. $G(D)$ is not always the unique element of $\mathbb{T}_D$ which reaches the maximum number of fixed length closed walks: for instance, for any $T \in \mathbb{T}_D$, we have $C(2;T) = 2|E(T)|$, which only depends on D.

Theorem 8.10 *Let D be a degree sequence of a tree. For any element $T \in \mathbb{T}_D$ and any $k \geq 0$, we have*

$$\mathrm{M}_k(T) = C(k;T) \leq C(k;G(D)) = \mathrm{M}_k(G(D)).$$

Moreover, the inequality is strict for sufficiently large even k if T and $G(D)$ are not isomorphic.

Proof. If it is possible to choose an edge or a vertex as root such that T is not level greedy, then we let T_1 be the level greedy tree with the same leveled degree sequence as T. We iterate this process: if an edge or vertex root can be chosen such that T_l is not level greedy, replace it by the corresponding level greedy tree, which we denote by T_{l+1}. Then $\mathrm{M}_k(T_{l+1}) \geq \mathrm{M}_k(T_l)$ for all $k \geq 0$, and for sufficiently large even k, the inequality is strict. Therefore, no infinite loops are possible in this process.

Hence there exists an integer m such that T_m is level greedy with respect to any choice of vertex or edge root. This tree T_m satisfies the "semi-regular" property defined in [78] (see Remark 7.6), and hence it is a greedy tree. From Theorems 8.4 and 8.9, we obtain

$$C(k;T) \leq C(k;T_1) \leq \cdots \leq C(k;T_m) = C(k;G(D))$$

for any $k \geq 0$, with strict inequality for sufficiently large even k. □

Remark 8.11 While the inequality in Theorem 8.10 is strict for sufficiently large k, there is no "universal" k with this property: for every k, there exists some degree sequence D and a tree T with degree sequence D that is not isomorphic to the greedy tree $G = G(D)$ such that

$$\mathrm{M}_\ell(T) = \mathrm{M}_\ell(G), \qquad \ell = 0, 1, \ldots, k.$$

Consider, for instance, the degree sequence $D = (3, 3, 2, 2, \ldots, 2, 1, 1, 1, 1)$, where the number of 2s is $4r - 2$ for some integer $r \geq 1$. The greedy tree $G = G(D)$ consists of two neighboring vertices of degree 3 to which paths are attached: two paths of length r to one of the two, two paths of length $r+1$ to the other. Now let T be the tree where one of the paths of length r in G is interchanged with one of the paths of length $r+1$.

T and G have the same number of (closed) walks of any length that do not contain the vertices of degree 3, since the forests resulting when the two are removed are isomorphic. Moreover, the subtrees of T and G consisting of vertices whose distance from the degree 3 vertices is at most r are isomorphic as well. Thus

$$\mathrm{M}_\ell(T) = \mathrm{M}_\ell(G), \qquad \ell \le 2r.$$

8.1.4 Consequences of the main result

Several corollaries follow immediately from our main theorem. In particular, in view of (8.5), we obtain the following corollary:

Corollary 8.12 *For any function $f(x) = \sum_{k=0}^{\infty} a_k x^k$ with nonnegative coefficients and for any tree T with degree sequence D, we have*

$$\mathrm{E}_f(T) \le \mathrm{E}_f(G(D)),$$

where E_f is defined as in (8.3)*. If the even part of f is not a polynomial (i.e., $a_k > 0$ for infinitely many even values of k), then the inequality is strict unless T is isomorphic to $G(D)$. In particular,*

$$\mathrm{EE}(T) < \mathrm{EE}(G(D))$$

for all $T \in \mathbb{T}_D$ that are not isomorphic to $G(D)$.

Moreover, we also obtain one of the main results of [10] as another corollary, since the spectral radius $\rho(T)$ of a tree T is equal to the limit $\lim_{\ell\to\infty} \sqrt[2\ell]{\mathrm{M}_{2\ell}(T)}$.

Corollary 8.13 *Among all trees with degree sequence D, the greedy tree $G(D)$ has the largest spectral radius $\rho(G(D))$.*

In [10], it was also shown that the greedy tree is unique with this property.

The Estrada index is just one of in principle infinitely many graph invariants of the form E_f. One could certainly conceive of a "Hyper-Estrada index", for example:

$$\mathrm{EEE}(G) = \sum_{i=1}^{n} e^{e^{\lambda_i}}.$$

A somewhat more natural example is the following: note that the characteristic polynomial of a graph G is given by

$$P_G(x) = \prod_{i=1}^{n}(x-\lambda_i) = x^n \prod_{i=1}^{n}\left(1 - \frac{\lambda_i}{x}\right).$$

If x is greater than the spectral radius, then we can take the logarithm and expand it into a power series:

$$\begin{aligned}\log P_G(x) &= n\log x + \sum_{i=1}^{n} \log\left(1 - \frac{\lambda_i}{x}\right) = n\log x - \sum_{i=1}^{n}\sum_{k=1}^{\infty} \frac{1}{k}\cdot\frac{\lambda_i^k}{x^k} \\ &= n\log x - \sum_{k=1}^{\infty} \frac{M_k(G)}{kx^k}.\end{aligned}$$

This formula, together with our main result, implies the following statement:

Corollary 8.14 *For any tree T with degree sequence D and any $x > \rho(G(D))$, the inequality*

$$P_T(x) \geq P_{G(D)}(x)$$

holds, with equality only if T is isomorphic to $G(D)$.

8.2 Trees with different degree sequences

In this section, we compare greedy trees with different degree sequences. This allows us to determine the maximal spectral moments of trees with different restrictions, e.g. given maximum degree or number of leaves.

To this end, we use a transformation on level greedy trees, where branches are moved between vertices at the same level. We study the effect of such a transformation on the number of closed walks of given length. Unlike the procedure in the proof of Theorem 8.10, the transformation that we consider in the following lemma does not preserve the degree sequence.

For any vertex v in a rooted tree T, we denote by T_v the rooted tree spanned by v and all its descendants, where v is chosen to be the root.

Lemma 8.15 *Let $D = ((i_{1,1}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$ be a leveled degree sequence of a (vertex) rooted tree. For some i and j with $1 < i < L(D)$ and $1 < j \leq k_i$, let B be a branch of g_j^i in the level greedy tree $G = G(D)$ which does not contain the root. Choose the neighbor of g_j^i in B to be the root of B. Let $T = G - g_j^i\,\mathrm{r}(B) + g_{j'}^i\,\mathrm{r}(B)$ for some $j' < j$ (see Figure 8.3). Then we have*

$$C(k;T) \geq C(k;G)$$

for any nonnegative integer k. For even $k \geq 4$, the inequality is strict.

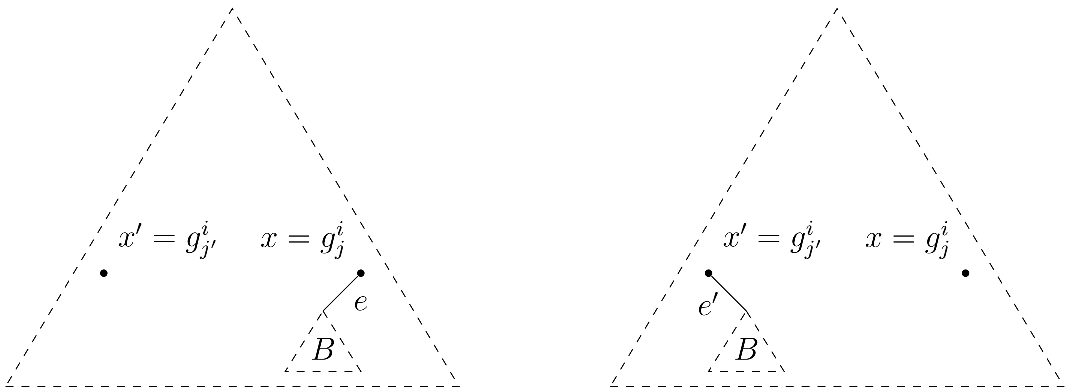


Figure 8.3: Moving a branch: the level greedy tree G (left) and the resulting tree T (right)

Proof. We use the same labels for vertices in T as in G. For notational convenience, set $x = g^i_j$, $x' = g^i_{j'}$, $e = g^i_j\,\mathrm{r}(B)$ and $e' = g^i_{j'}\,\mathrm{r}(B)$.

It is clear that $C(k;T) - C^{e'}(k;T) = C(k;G) - C^{e}(k;G)$ because $T - e' = G - e$. Thus it suffices to prove

$$C^{e'}(k;T) \geq C^{e}(k;G). \tag{8.27}$$

Let $v = g^{i'}_l$ be the closest common ancestor of $x = g^i_j$ and $x' = g^i_{j'}$ in G, and let $u = g^{i'+1}_h$ and $u' = g^{i'+1}_{h'}$ be the neighbors of v in the branches containing g^i_j and $g^i_{j'}$ respectively.

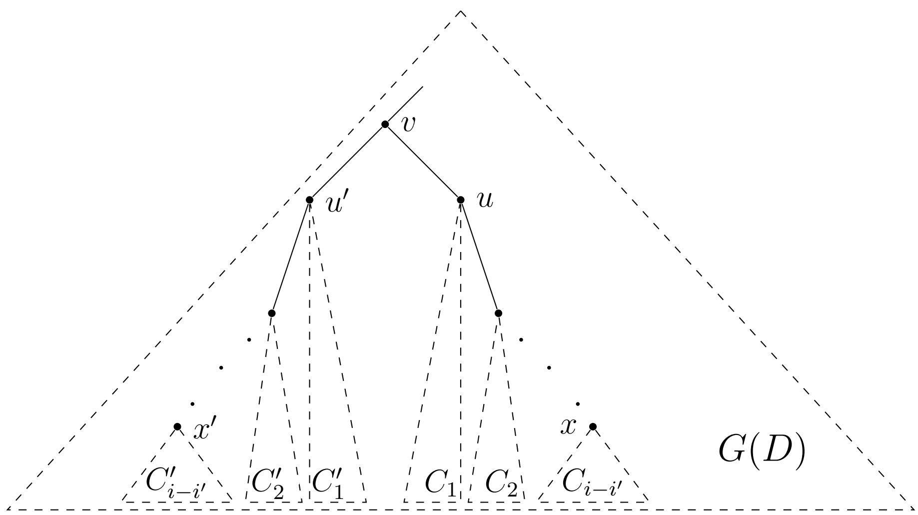


Figure 8.4: Decomposition of G_v in the proof of Lemma 8.15

Consider a decomposition of G as in Figure 8.4. Let $P = vv_1 \ldots v_{i-i'}$ and $P' = vv'_1 \ldots v'_{i-i'}$ be the paths joining v to $v_{i-i'} := x$ and $v'_{i-i'} := x'$, respectively. For any $1 \leq t \leq i - i'$, we define C_t and C'_t to be the largest branch in G which contains v_t and v'_t respectively such that $V(C_t) \cap V(P) = \{v_t\}$ and $V(C'_t) \cap V(P') = \{v'_t\}$. Since G is level greedy, there is an isomorphism preserving roots between C_r and a subgraph of C'_r for any $r \in \{1, 2, \ldots, i - i'\}$. Therefore one can find an injective homomorphism, say $f : V\big(G_u\big) \longrightarrow V\big(T_{u'}\big)$, which satisfies $f(u) = u'$, $f(x) = x'$ and $f(e) = e'$.

The map

$$F : \mathcal{C}^e(k; G) - \mathcal{C}^{v,e}(k; G) \longrightarrow \mathcal{C}^{e'}(k; T) - \mathcal{C}^{v,e'}(k; T)$$
$$w_1 \ldots w_{k+1} \longmapsto f(w_1) \ldots f(w_{k+1})$$

is injective because f is injective. We also define a map

$$F' : \mathcal{C}^v(k; G) \longrightarrow \mathcal{C}^v(k; T)$$

in a recursive way. Let $W = w_1 \ldots w_{k+1} \in \mathcal{C}^v(k; G)$, and let m and M be, respectively, the smallest and largest integers such that $w_m = w_M = v$ and $1 < m \leq M < k + 1$, if there exist such integers. Then we define:

- If $v \notin \{w_2, \ldots, w_k\}$ (and hence $w_1 = w_{k+1} = v$) and $w_s w_{s+1} \neq e$ for any $s = 1, \ldots, k$, then $F'(W) = w_1 \ldots w_{k+1}$.

- If $v \notin \{w_2, \ldots, w_k\}$ and $w_s w_{s+1} = e$ for some $s \in \{1, \ldots, k\}$, then

$$F'(W) = w_1 f(w_2) \ldots f(w_k) w_{k+1}.$$

- Otherwise we set

$$F'(W) = \phi(w_1 \ldots w_{m-1}) F'(w_m \ldots w_M) \phi(w_{M+1} \ldots w_{k+1}),$$

  where $\phi(w_1 \ldots w_{m-1}) = f(w_1) \ldots f(w_{m-1})$ if $w_s w_{s+1} = e$ for some $s \in \{1, \ldots, m - 2\}$, and $\phi(w_1 \ldots w_{m-1}) = w_1 \ldots w_{m-1}$ otherwise.

In words, we break a walk into pieces separated by visits to vertex v. Each piece is either kept the same (if it does not contain e) or replaced by its image under the injection f if it contains e. Since the decomposition is unique and f is injective, the so constructed map F' is also an injection, and so is its restriction to $\mathcal{C}^{v,e}(k; G)$. This proves inequality (8.27) and thus the main inequality.

For even $k \geq 4$, the inequality is strict, since F is not surjective. The degree of x in G is strictly less than the degree of x' in T by construction. Hence, there is an edge e'' incident to x' that does not have a preimage under F, and so is any walk starting from e' that uses e''. There is such a closed walk for arbitrary even length larger than 4. $\square$

Lemma 8.16 *Let $D = ((i_{1,1}, i_{1,2}), (i_{2,1}, \dots, i_{2,k_2}), \dots, (i_{n,1}, \dots, i_{n,k_n}))$ be a leveled degree sequence of an edge-rooted tree. For some i and j with $1 \le i < L(D)$ and $1 < j \le k_i$, let B be a branch of g^i_j in the level greedy tree $G = G(D)$ which does not contain the root. Take the neighbor of g^i_j in B as root of B. Let $T = G - g^i_j\,\mathrm{r}(B) + g^i_{j'}\,\mathrm{r}(B)$ for some $j' < j$. Then we have*

$$C(k;T) \ge C(k;G)$$

for any nonnegative integer k. For even $k \ge 4$, the inequality is strict.

Proof. Again, we keep the labels of vertices of G in T. For simplicity, we write $x = g^i_j$, $x' = g^i_{j'}$, $e = g^i_j\,\mathrm{r}(B)$, $e' = g^i_{j'}\,\mathrm{r}(B)$, $r = \mathrm{r}(G)$ and $r' = \mathrm{r}(T)$.

Let G_1 and G_2 be the components of $G - r$, and T_1 and T_2 those of $T - r'$, such that $|V(G_1)| \ge |V(G_2)|$ and $|V(T_1)| \ge |V(T_2)|$.

If x and x' are both vertices of the same component G_m, then the proof is exactly the same as that of Lemma 8.15. So from now on, we assume that $x \in V(G_2)$ and $x' \in V(G_1)$.

If G is decomposed as in Figure 8.5, then C_r has a copy preserving levels in C'_r for any $1 \le r \le i$. Because of this fact, we know that one can find a level preserving injective homomorphism, say f, between G_2 and T_1 (which has G_1 as a subgraph) which satisfies $f(g^1_2) = g^1_1$, $f(x) = x'$ and $f(e) = e'$.

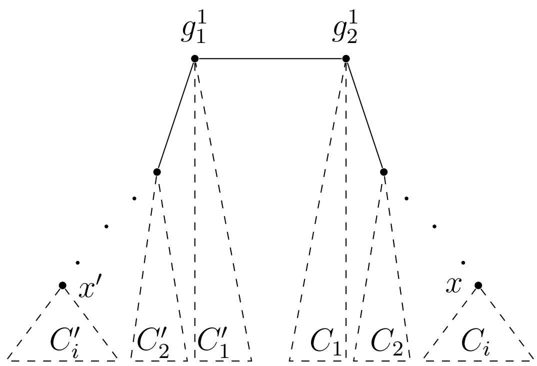


Figure 8.5: Decomposition of G in the proof of Lemma 8.16

Since we deal with closed walks, we are only interested in even $k = 2l$. We know that

$$C(2l;T) - C^{r'}(2l;T) - [C(2l;G) - C^{r}(2l;G)] = C^{e'}(2l;T_1) - C^{e}(2l;G_2)$$

is nonnegative: as f is injective, so is the map

$$\begin{aligned} F : \mathcal{C}^{e}(2l;G_2) &\longrightarrow \mathcal{C}^{e'}(2l;T_1) \\ w_1 \dots w_{k+1} &\longmapsto f(w_1) \dots f(w_{k+1}). \end{aligned}$$

Since f is level preserving, we can even choose an arbitrary level sequence S of walks without two consecutive 1s and still have

$$\begin{aligned} C^{e'}(S;T_1) &\geq C^e(S;G_2),\\ C(S,(T_1-B)\cup T_2) &= C(S;G_1\cup(G_2-B)), \end{aligned} \tag{8.28}$$

and hence $C(S;T_1\cup T_2)\geq C(S;G_1\cup G_2)$. Now we are left to show that

$$C^{r'}(2l;T)-C^{r}(2l;G)=C^{e',r'}(2l;T)-C^{e,r}(2l;G)\geq 0 \tag{8.29}$$

for any integer $l\geq 1$. Before that let us first show that

$$C^{e'}_{r'}(2l;T)\geq C^{e}_{r}(2l;G) \tag{8.30}$$

for any positive integer l. Note the subtle difference between $C^{e,r}(2l;G)$ and $C^e_r(2l;G)$ here: the former counts walks that pass through r *at some stage*, while the latter counts walks that start with r. We reason by induction on l. For $l=1$ we have $C^{e'}_{r'}(2;T)=C^e_r(2;G)=0$. Assume that (8.30) holds whenever $l\leq m$ for some $m\geq 1$. Since

$$C_{r'}(2l;T)-C^{e'}_{r'}(2l;T)=C_{r'}(2l;T-B)=C_r(2l;G-B)=C_r(2l;G)-C^e_r(2l;G)$$

for all l, the induction hypothesis also implies that $C_{r'}(2l;T)\geq C_r(2l;G)$ for all $l\leq m$.

Consider now the case where $l=m+1$. Let

$$\mathcal{C}^{e'}_{r'}(2l;T)=P_1(l)\cup Q_1(l)\cup R_1(l)\qquad\text{and}\qquad \mathcal{C}^{e}_{r}(2l;G)=P_2(l)\cup Q_2(l)\cup R_2(l),$$

where the $P_i(l)$'s contain walks whose level sequences start with $1,1,1,1$, the $Q_i(l)$'s contain walks whose level sequences start with $1,1,1,2$, and the level sequences of the elements of the $R_i(l)$'s start with $1,1,2$. The induction hypothesis implies

$$|P_1(l)|=C^{e'}_{r'}(2(l-1);T)\geq C^e_r(2(l-1);G)=|P_2(l)|.$$

It is easy to check that $|Q_1(1)|=|Q_2(1)|=0$, $|Q_1(2)|=|Q_2(2)|\in\{0,1\}$. For $l\geq 3$, we define for any j with $2\leq j\leq 2l-3$ the subset $Q_i^j(l)$ of $Q_i(l)$ whose elements have level sequence $(1,1,1,i_1,\ldots,i_{2l-2})$, where $i_j=i_{j+1}=1$ and $(i_s,i_{s+1})\neq(1,1)$ for $s=1,\ldots,j-1$. It is convenient to also define

$$Q_i^{2l-2}(l)=Q_i(l)-\bigcup_{j=2}^{2l-3}Q_i^j(l),$$

it contains the elements of $Q_i(l)$ with level sequence $(1,1,1,i_1,\ldots,i_{2l-2})$, where $i_1=2$ and $(1,1)\neq(i_s,i_{s+1})$ for $s=1,\ldots,2l-3$. Set

$$\mathbb{S}^1_j=\{(1,i_1,\ldots,i_{j-1},1): i_1=i_{j-1}=2,(1,1)\neq(i_s,i_{s+1})\text{ for } s\in\{1,\ldots,j-2\}\}.$$

Now we decompose walks in Q_1^j and Q_2^j: any such walk consists of two steps along the edge root (forwards and backwards), then continues to higher levels and returns to the first level after j steps (possibly earlier as well, but without ever using the edge root). We call this part U_1; its level sequence lies in $\mathbb{S}_j^1$. Thereafter, the walk continues for another $2l-j-2$ steps, starting with the edge root; this part is called U_2. Since we know that a walk in Q_1^j has to pass through e', we have the following possibilities:

- The walk U_1 uses e' (which means that it lies entirely in T_1), the walk U_2 is arbitrary.
- The walk U_1 does not use e', but stays in T_1 (thus it lies in T_1-B), the walk U_2 contains e'.
- The walk U_1 lies in T_2, thus it does not use e'. Then the walk U_2 has to contain e'.

For Q_2^j, there are three analogous possibilities. Making use of this decomposition, Lemma 8.6, (8.28) and the induction hypothesis, we obtain

$$\begin{aligned}
|Q_1^j(l)| &= \sum_{S\in\mathbb{S}_j^1} C^{e'}(S;T_1)C_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(2l-j-2;T)\\
&\quad+\sum_{S\in\mathbb{S}_j^1} C(S;T_1-B)C^{e'}_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(2l-j-2;T)\\
&\quad+\sum_{S\in\mathbb{S}_j^1} C(S;T_2)C^{e'}_{\mathrm{r}(T_2),\mathrm{r}(T_1)}(2l-j-2;T)\\
&= \sum_{S\in\mathbb{S}_j^1} C^{e'}(S;T_1)\cdot\frac{1}{2}C_{r'}(2l-j-2;T)\\
&\quad+\sum_{S\in\mathbb{S}_j^1} \left(C(S;T_1-B)+C(S;T_2)\right)\cdot\frac{1}{2}C^{e'}_{r'}(2l-j-2;T)\\
&\geq \sum_{S\in\mathbb{S}_j^1} C^{e}(S;G_2)\cdot\frac{1}{2}C_{r}(2l-j-2;G)\\
&\quad+\sum_{S\in\mathbb{S}_j^1} \left(C(S;G_1)+C(S;G_2-B)\right)\cdot\frac{1}{2}C^{e}_{r}(2l-j-2;G) = |Q_2^j(l)|
\end{aligned}$$

for all j such that $2\leq j\leq 2l-3$. For $j=2l-2$, walk U_2 is empty, so we have

$$|Q_1^{2l-2}(l)| = \sum_{S\in\mathbb{S}_j^1} C^{e'}(S;T_1) \geq \sum_{S\in\mathbb{S}_j^1} C^{e}(S;G_2) = |Q_2^{2l-2}(l)|.$$

We conclude with the third subclass of walks whose level sequences start with $1,1,2$. For any j with $2 \le j \le 2l-2$, let $R_i^j(l)$ be the subset of $R_i(l)$ whose elements have level sequence $(1,1,i_1,\dots,i_{2l-1})$, where $i_j = 1$ and $1 \notin \{i_1,\dots,i_{j-1}\}$. The case that $j = 2l-1$ is not interesting since it does not correspond to any closed walk. We decompose $R_1^j(l)$ and $R_2^j(l)$ in a similar way as we decomposed $Q_1^j(l)$ and $Q_2^j(l)$. Define

$$\mathbb{S}_j^2 = \{(1,i_1,\dots,i_{j-1},1) : 1 \notin \{i_1,\dots,i_{j-1}\}\}.$$

A walk in $R_1^j(l)$ (or $R_2^j(l)$) consists of a step along the edge root, then moves to higher levels and only returns to the first level after j steps. This part of j steps has a level sequence in $\mathbb{S}_j^2$, the rest forms a closed walk starting with the edge root. Dividing into three cases again, depending on which part contains e' (e, respectively), we obtain

$$\begin{aligned}
|R_1^j(l)| &= \sum_{S\in\mathbb{S}_j^2} C^{e'}(S;T_1)C_{\mathrm{r}(T_2),\mathrm{r}(T_1)}(2l-j;T) \\
&\quad + \sum_{S\in\mathbb{S}_j^2} C(S;T_1-B)C^{e'}_{\mathrm{r}(T_2),\mathrm{r}(T_1)}(2l-j;T) \\
&\quad + \sum_{S\in\mathbb{S}_j^2} C(S;T_2)C^{e'}_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(2l-j;T) \\
&= \sum_{S\in\mathbb{S}_j^2} C^{e'}(S;T_1)\cdot\frac{1}{2}C_{r'}(2l-j;T) \\
&\quad + \sum_{S\in\mathbb{S}_j^2} (C(S;T_1-B)+C(S;T_2))\cdot\frac{1}{2}C^{e'}_{r'}(2l-j;T) \\
&\ge \sum_{S\in\mathbb{S}_j^2} C^{e}(S;G_2)\cdot\frac{1}{2}C_{r}(2l-j;G) \\
&\quad + \sum_{S\in\mathbb{S}_j^2} (C(S;G_1)+C(S;G_2-B))\cdot\frac{1}{2}C^{e}_{r}(2l-j;G) = |R_2^j(l)|.
\end{aligned}$$

This completes the proof of (8.30). We now proceed to the proof of (8.29), making use of a similar argument as in Lemma 8.8. Any element, say W, in $\mathcal{C}^{e',r'}(2l;T)$ or $\mathcal{C}^{e,r}(2l;G)$ has a unique decomposition

$$W = W_1W_2W_3, \tag{8.31}$$

where W_2 is a closed walk that starts with the edge root and is chosen to have maximal length and $W' = W_1W_3$ forms a closed walk which never uses the edge root, but passes at least once through one of its ends (unless it is empty). The decomposition (8.31) is unique (as it was explained in the proof of Lemma 8.8). Now let

$$\mathbb{S}_j^3 = \{(i_1,\dots,i_{j+1}) : i_s = 1 \text{ for some } 1 \le s \le j+1\}.$$

The walk W' has a level sequence in $\mathbb{S}_j^3$ for some j. Again, there are three possibilities for a walk in $\mathcal{C}^{e',r'}(2l;T)$:

- The walk W' contains e', and thus lies entirely in T_1, and W_2 is arbitrary.
- The walk W' does not contain e', but still lies in T_1 (thus entirely in $T_1 - B$), and W_2 uses e'.
- The walk W' lies in T_2, thus does not use e'. Then W_2 has to use e'.

There are three analogous possibilities for $\mathcal{C}^{e,r}(2l;G)$. We obtain

$$\begin{aligned}
C^{e',r'}(2l;T) &= \sum_{j=0}^{2l-2}\sum_{S\in\mathbb{S}_j^3} C^{e'}(S;T_1)C_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(2l-j;T)\\
&\quad + C(S;T_1-B)C^{e'}_{\mathrm{r}(T_1),\mathrm{r}(T_2)}(2l-j;T) + C(S;T_2)C^{e'}_{\mathrm{r}(T_2),\mathrm{r}(T_1)}(2l-j;T)\\
&= \sum_{j=0}^{2l-2}\sum_{S\in\mathbb{S}_j^3} C^{e'}(S;T_1)\cdot\frac{1}{2}C_{r'}(2l-j;T)\\
&\qquad + (C(S;T_1-B)+C(S;T_2))\cdot\frac{1}{2}C^{e'}_{r'}(2l-j;T)\\
&\geq \sum_{j=0}^{2l-2}\sum_{S\in\mathbb{S}_j^3} C^{e}(S;G_2)\cdot\frac{1}{2}C_{r}(2l-j;G)\\
&\qquad + (C(S;G_1)+C(S;G_2-B))\cdot\frac{1}{2}C^{e}_{r}(2l-j;G) = C^{e,r'}(2l;G).
\end{aligned}$$

This concludes the proof of (8.29) and thus the theorem. As in the previous lemma, the inequality is strict for even $k \geq 4$ since the map F is not surjective.

□

Given two degree sequences $B \preccurlyeq D$ of trees, by iteratively transferring branches, we can transform $G(B)$ to become an element of $\mathbb{T}_D$. As seen in the proof of the next theorem, it turns out that it is always enough to only use transfers of the type described in the two Lemmas 8.15 and 8.16 to obtain an element of $\mathbb{T}_D$ from $G(B)$, showing that $G(D)$ has more closed walks of any length than $G(B)$. This parallels analogous results for e.g. the number of subtrees [94] or the spectral radius [10].

Theorem 8.17 *Let $D = (d_1,\dots,d_n)$ and $B = (b_1,\dots,b_n)$ be degree sequences of trees of the same order such that $B \preccurlyeq D$. Then for any integer $k \geq 0$ we have*

$$C(k;G(B)) \leq C(k;G(D)).$$

If $B \neq D$ and k is even and ≥ 4, then the inequality is strict.

Proof. The statement is obvious for $B = D$. From now, we assume that there exists some i_0 such that $b_{i_0} \neq d_{i_0}$. Since

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} d_i, \tag{8.32}$$

we know that the set $\{i : d_i \neq b_i\}$ must have at least two elements. Let $l = \min\{i : d_i \neq b_i\}$ and $m = \max\{i : d_i \neq b_i\}$. We must have $b_l < d_l$, $b_m > d_m$ and hence $b_{l-1} = d_{l-1} \geq d_l \geq b_l + 1$ and $b_{m+1} = d_{m+1} \leq d_m \leq b_m - 1$. Therefore, $B_1 = (b_1, \ldots, b_{l-1}, b_l + 1, b_{l+1}, \ldots, b_{m-1}, b_m - 1, b_{m+1}, \ldots, b_n)$ is a valid degree sequence. It is easy to see that $B \preccurlyeq B_1$. Consider two vertices u and v in the greedy tree $G(B)$ such that $\deg u = b_l$ and $\deg v = b_m$.

Case 1: The length of the path in $G(B)$ joining u and v is even. Let w be the middle vertex of this path. Consider $G(B)$ as a level greedy tree whose root is w. Then u and v are on the same level, say level h. We have $u = g_i^h$ and $v = g_j^h$ for some $i < j$. Let $w = g_r^{h+1}$ be a child of $v = g_j^h$, and let $H = G(B)_w$ be the branch rooted at w.

Consider $T = G(B) - vw + uw$; the degree sequence of T is B_1. By Theorem 8.4 and Lemma 8.15, it follows that

$$C(k; G(B_1)) \geq C(k; T) \geq C(k; G(B))$$

for all $k \geq 0$.

Case 2: The length of the path in $G(B)$ joining u and v is odd. The argument is analogous to the previous case: we choose the middle edge of the path as root and then we use Theorem 8.9 and Lemma 8.16 instead of Theorem 8.4 and Lemma 8.15.

In either case, we have

$$C(k; G(B_1)) \geq C(k; G(B))$$

for all $k \geq 0$. We repeat this process to obtain a sequence of degree sequences $B_0 = B, B_1, B_2, \ldots, B_r = D$ such that $B = B_0 \preccurlyeq B_1 \preccurlyeq \cdots \preccurlyeq B_r = D$ and

$$C(k; G(B)) = C(k; G(B_0)) \leq C(k; G(B_1)) \leq \cdots \leq C(k; G(B_r)) = C(k; G(D))$$

for all $k \geq 0$, which proves the theorem. $\square$

Conjecture 8.1 follows as corollary of the two Theorems 8.10 and 8.17: The degree sequence of the n-vertex Volkmann tree, which is of the form $(\Delta, \ldots, \Delta, r, 1, \ldots, 1)$ for some $1 \leq r < \Delta$, majorizes all possible degree sequences of n-vertex trees with maximum degree Δ.

More results can be obtained by similar arguments in the same way as Corollaries 5.1 – 5.5 of [94] and Corollaries 29 – 32 of [4] are obtained. Let us state some more of these corollaries, which also recover some results that can be found in [17, 89]:

Corollary 8.18 *For any n-vertex tree T and for any $k \geq 0$,*

$$\mathrm{M}_k(S_n) \geq \mathrm{M}_k(T),$$

where S_n is the star with n vertices, whose degree sequence is $(n-1, 1, \ldots, 1)$.

Corollary 8.19 *Among trees T of order n with s leaves, $\mathrm{M}_k(T)$ is maximized by the greedy tree $G(s, 2, 2, \ldots, 2, 1, 1, \ldots, 1)$ (the number of 2s is $n-s-1$, the number of 1s is s) for any $k \geq 0$.*

Corollary 8.20 *Among trees T of order n with independence number $\alpha \geq n/2$ and among all trees T with matching number $n - \alpha \leq n/2$, $\mathrm{M}_k(T)$ is maximized by the greedy tree $G(\alpha, 2, 2, \ldots, 2, 1, 1, \ldots, 1)$ (the number of 2s is $n - \alpha - 1$, the number of 1s is α) for any $k \geq 0$.*

M_k in each of the above corollaries can of course be replaced by EE or more generally E_f for any f with nonnegative coefficients in (8.4). If infinitely many even-indexed coefficients are strictly positive (e.g., for EE), then we even have strict inequality. Moreover, corollaries analogous to Corollary 8.13 and Corollary 8.14 for the spectral radius and the values of the characteristic polynomial also follow easily.

Chapter 9

Number of subtrees of trees with given degree sequence

The number of subtrees of a tree, being an interesting topic on its own mathematical right [77,87], also plays a role in other fields such as phylogenetic reconstruction [63]. In many questions concerning subtrees of trees, the number of subtrees of a specific order plays an important role. For instance, the average subtree order [59, 60, 81] and the subtree poset [58, 79, 80] have been considered in recent works. The concept of a subtree polynomial akin to the matching polynomial, the independence polynomial and other polynomials associated to a graph, has been brought forward as well [59]: if $n_k(T)$ is the number of subtrees of order k in a tree T of order n, then the associated polynomial is

$$\Phi_T(x) = \sum_{k=0}^{n} n_k(T)x^k.$$

More generally, a weighted version (also including edge weights) is studied in [87], and a bivariate version, where a second variable marks the number of leaves, is considered in [70].

In this chapter, we study n_k in the class of trees with given degree sequence.

9.1 Statement of results and preliminaries

The first main theorem of this chapter is a stronger version of that of [94]: the greedy tree maximizes the number of subtrees of any given order.

Theorem 9.1 *Among all trees T with degree sequence D, the number $n_k(T)$ of subtrees of order k attains its maximum when T is the greedy* *tree* $G(D)$*.*

It should be remarked that for specific k, the greedy tree is not necessarily the only tree for which $n_k(T)$ reaches its maximum (for instance, when $k = 1$, then $n_k(T)$ is simply the order of T and thus equal for all trees with degree sequence D). The important point of Theorem 9.1 is the fact that $n_k(T) \le n_k(G(D))$ for *all* k simultaneously and all trees T with degree sequence D.

Theorem 9.1 will be proven in Section 9.2. In fact, a slightly more general result holds:

Theorem 9.2 *Among all forests* F *with given degree sequence, the number* $n_k(F)$ *of subtrees of order* k *attains its maximum when* F *is the greedy forest.*

In Section 9.3, we compare greedy trees with different degree sequences, which yields a number of corollaries such as:

Corollary 9.3 *Among trees with given order* n *and maximum degree* Δ*, the number* $n_k(T)$ *of subtrees of order* k *attains its maximum when* T *is the greedy tree* $G(\Delta, \Delta, \ldots, \Delta, d, 1, 1, \ldots, 1)$*, where* $1 \le d < \Delta$ *is chosen in such a way that* $d \equiv n - 1 \bmod (\Delta - 1)$*.*

Similar results are obtained for trees with given number of leaves, independence number or matching number, as in the previous chapter.

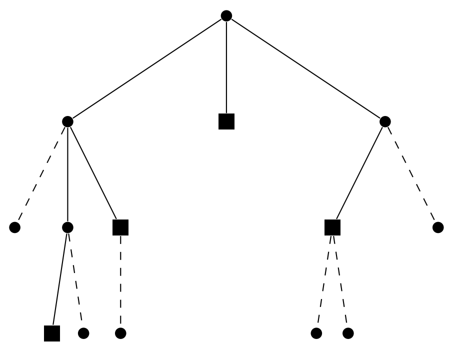


Figure 9.1: The correspondence between antichains (indicated by square nodes) and subtrees that contain the root (solid edges)

In Section 9.4, we study the number of subtrees containing a specific vertex (which we can assume to be the root). One of the motivations is a connection to a different counting problem: a rooted tree can be regarded as the Hasse diagram of a poset. There is a natural bijection between antichains and subtrees containing the root (see Figure 9.1): to each subtree that contains the root, we can associate the antichain that is formed by the leaves (excluding the root unless it is the only vertex of the subtree).

Therefore, the total number of subtrees that contain the root is the same as the number of (nonempty) antichains in the associated poset. It was pointed out by Klazar [61] that the number of antichains in a rooted tree of order n is at most $2^{n-1}+1$ (with equality when the tree is a star rooted at its center) and at least n (with equality when the tree is a path rooted at one of its ends). Apart from this, not much seems to be known about extremal values of the number of antichains in a rooted tree.

The main result of Section 9.4 reads as follows:

Theorem 9.4 *Let $n_k(T,v)$ denote the number of subtrees of order k in T that contain the vertex v. For any tree T with degree sequence D, any vertex v of T and any $k \geq 1$, the inequality*

$$n_k(T,v) \leq n_k(G(D), r(G(D)))$$

holds, where $r(G(D))$ is the canonical root of the greedy tree, as chosen in Definition 7.5.

This implies that the greedy tree, rooted in the canonical way, also has the greatest number of antichains among all rooted trees with given degree sequence. A more general statement, where the degree of the root can be prescribed as well, also holds (see Theorem 9.22). Moreover, we also prove analogous statements for subtrees with a given number of leaves, which corresponds to antichains of given cardinality.

Now we introduce further notation and state a few remarks.

We denote by $\mathbb{F}_D$ the set of all rooted forests with two components and with given leveled degree sequence D.

Let T_1 and T_2 be two rooted trees. For $j \in \{1,2\}$ and $l \geq 1$ let $\mathcal{V}_{l,j} = \{v^l_{j,1}, \ldots, v^l_{j,k_{l,j}}\}$ be the set of vertices at the l^{th} level of T_j. We write $T_1 \rhd T_2$ if the height of T_1 is at least that of T_2 and for any $l \geq 1$ we have

$$\min\left\{\deg v^l_{1,1}, \ldots, \deg v^l_{1,k_{l,1}}\right\} \geq \max\left\{\deg v^l_{2,1}, \ldots, \deg v^l_{j,k_{l,2}}\right\}$$

if $\mathcal{V}_{l,2}$ is not empty. The relation $\rhd$ is easily seen to be transitive.

Remark 9.5 Let F be a rooted forest. F is a level greedy forest if and only if its components can be labeled as $F_1, \ldots, F_t$ such that each of $F_1, \ldots, F_t$ is a level greedy tree and $F_1 \rhd \cdots \rhd F_t$. A tree T rooted at v is a level greedy tree if and only if $T - v$ is a level greedy forest.

The following simple observation turns out to be extremely useful in this study.

Lemma 9.6 *Let F be a rooted forest with $t \geq 2$ components. F is a level greedy forest if and only if any two components of F form a level greedy forest.*

Proof. As mentioned in Remark 9.5, a rooted forest is a level greedy forest if and only if $\triangleright$ induces a total order on its components (it is possible that two components F_i and F_j are isomorphic, but this can only happen if their degrees are constant on each level, in which case $F_i \triangleright F_j$ holds as well as $F_j \triangleright F_i$). Equivalently, any two components should be comparable by $\triangleright$, which in turn is equivalent to the statement that any two components form a level greedy forest. $\square$

9.2 The number of subtrees of given order

We prove in this section that, given the degree sequence, the greedy tree maximizes the number of subtrees of any order k. We first consider two-component forests with a given leveled degree sequence, the case of a tree can then be considered as a special case.

We denote by $G_1(D)$ and $G_2(D)$ the two connected components of the level greedy forest $G(D)$, where we assume that

$$|V(G_1(D))| \geq |V(G_2(D))|.$$

Similarly, we write F_1 and F_2 for the components of a two-component rooted forest. In order to formulate our key lemma (Lemma 9.8 below), we need a few more definitions:

Definition 9.7 Let F be a rooted forest which has n levels of vertices. The *level sequence* of a subforest F' of F is the sequence $(s_1, \ldots, s_n)$, where s_i is the number of vertices of F' at the i^{th} level in F.

We denote by $n_S(F)$ the number of subtrees in F with level sequence S. Clearly, for any integer $k \geq 1$, the number $n_k(F)$ of subtrees of order k in F can be written as the sum of $n_S(F)$ over all possible level sequences S that sum to k. Given a level sequence S, we write S^- for the level sequence obtained from S by removing the first term (i.e., if $S = (s_1, \ldots, s_n)$, then $S^- = (s_2, \ldots, s_n)$).

Now we are ready to formulate and prove the key lemma of this section.

Lemma 9.8 *Let D be a leveled degree sequence of a two-component forest. For any level sequence $S = (s_1, s_2, \ldots, s_{L(D)})$ and for any $F \in \mathbb{F}_D$ we have*

$$(n_S(F_1), n_S(F_2)) \blacktriangleleft (n_S(G_1(D)), n_S(G_2(D))). \tag{9.1}$$

Similarly, if D is a leveled degree sequence of a (vertex) rooted tree, then for any $T \in \mathbb{T}_D$ and for any level sequence $S = (s_1, s_2, \ldots, s_{L(D)})$ we have

$$n_S(T) \leq n_S(G(D)). \tag{9.2}$$

Remark 9.9 Note that (9.2) is equivalent to $(n_S(T), 0) \blacktriangleleft (n_S(G(D)), 0)$. Hence we will only show (9.1), where F_2 is allowed to be empty.

Proof. We prove (9.1) by induction on $L(D)$. The case of $L(D) = 1$, as well as that of $L(D) = 2$, is trivial: in either case, the corresponding sets $\mathbb{F}_D$ and $\mathbb{T}_D$ contain only one element. Assume that the lemma is true whenever $L(D) \leq k$ for some integer $k \geq 2$. Now assume that $L(D) = k+1$. We can also assume that $n_S(F_1) \geq n_S(F_2)$. There are two cases:
Case 1: $s_1 = 0$. In this case we have

$$n_S(G(D)) = \sum_{X \in C(G_1(D)) \cup C(G_2(D))} n_{S^-}(X)$$

and

$$n_S(F) = \sum_{X \in C(F_1) \cup C(F_2)} n_{S^-}(X).$$

Assume that there are elements H_1 and H_2 of $C(F_1) \cup C(F_2)$ such that $H_1 \cup H_2$ is not a (rooted) greedy forest, and let B be the leveled degree sequence of $H_1 \cup H_2$. We know by the induction hypothesis that

$$(n_{S^-}(H_1), n_{S^-}(H_2)) \blacktriangleleft (n_{S^-}(G_1(B)), n_{S^-}(G_2(B))).$$

By replacing H_1 and H_2 by $G_1(B)$ and $G_2(B)$, respectively, we obtain a new rooted forest F^1 with the same leveled degree sequence: if H_1 and H_2 both belong to $C(F_1)$, then $G_1(B)$ and $G_2(B)$ become part of $C(F_1^1)$, and the same applies to $C(F_2)$ and $C(F_2^1)$. If H_1 and H_2 belong to $C(F_1)$ and $C(F_2)$ respectively, then the larger component $G_1(B)$ becomes part of $C(F_1^1)$ and $G_2(B)$ becomes part of $C(F_2^1)$. It follows that

$$(n_S(F_1), n_S(F_2)) \blacktriangleleft (n_S(F_1^1), n_S(F_2^1)).$$

We iterate the process to obtain a sequence (with $F = F^0$)

$$(n_S(F_1^0), n_S(F_2^0)) \blacktriangleleft (n_S(F_1^1), n_S(F_2^1)) \blacktriangleleft \cdots \blacktriangleleft (n_S(F_1^K), n_S(F_2^K)).$$

This process always terminates, since the vector of all component sizes of $F^{t+1} - r(F^{t+1})$ (sorted in descending order) is lexicographically greater than that of $F^t - r(F^t)$. At the end, any two elements of $C(F_1^K) \cup C(F_2^K)$ form a greedy forest. By Lemma 9.6, such a situation is reached only when $F^K - \{r(F_1^K), r(F_2^K)\}$ is level greedy. Thus the branches of F^K are the same as those of $G(D)$, which means that

$$n_S(F_1^K) + n_S(F_2^K) = n_S(G_1(D)) + n_S(G_2(D)).$$

Since the largest branches are all part of $G_1(D)$ in the greedy tree, we also have

$$(n_S(F_1^K), n_S(F_2^K)) \blacktriangleleft (n_S(G_1(D), n_S(G_2(D))),$$

which completes the proof in this case.

Case 2: $s_1 = 1$. Let r_i be the degree of the root of F_i for $i = 1, 2$. We reason by induction on $r = \max\{r_1, r_2\}$. If $r = 1$ then there are two subcases:

- If $s_2 = 0$, there is nothing to prove: all potential subtrees only consist of a root, so that $n_S(F)$ does not actually depend on F. Likewise, if $s_2 \geq 2$, then there are no subtrees with level sequence S in view of the assumption $r = 1$, and this is independent of the shape of F.

- Assume that $s_2 = 1$. Let G_1', G_2', F_1', F_2' be the trees obtained by removing the roots from $G_1(D), G_2(D), F_1$, F_2 respectively (G_2' and F_2' are empty if $G_2(D)$ and F_2 are). We have

  $$n_S(G_1(D)) = n_{S^-}(G_1'),\ n_S(G_2(D)) = n_{S^-}(G_2'),$$

  and

  $$n_S(F_1) = n_{S^-}(F_1'),\ n_S(F_2) = n_{S^-}(F_2').$$

  Hence (9.1) follows (by our outer induction hypothesis with respect to $L(D)$) from

  $$(n_{S^-}(F_1'), n_{S^-}(F_2')) \blacktriangleleft (n_{S^-}(G_1'), n_{S^-}(G_2')).$$

Assume (9.1) holds for $r \leq l$ for some $l \geq 1$. Let $r = l + 1$, and let A, B, A', B' be subtrees of F_1 and F_2 as shown in Figure 9.2. Note that if F_2 is an isolated vertex then B is empty and if F_2 is empty then so are B and B'.

Let $\mathbb{S}$ be the set of all possible level sequences whose first term is 1, so that we have

$$\begin{aligned} n_S(F) &= \sum_{\substack{S_1, S_2 \in \mathbb{S} \\ S_1^- + S_2^- = S^-}} \big(n_{S_1}(A') n_{S_2}(F_1 - A) + n_{S_1}(B') n_{S_2}(F_2 - B)\big) \\ &= \sum_{\substack{S_1, S_2 \in \mathbb{S} \\ S_1^- + S_2^- = S^-}} \big(n_{S_1^-}(A) n_{S_2}(F_1 - A) + n_{S_1^-}(B) n_{S_2}(F_2 - B)\big). \end{aligned}$$

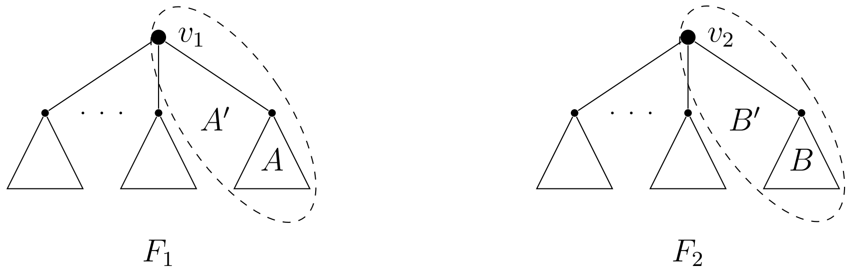


Figure 9.2: Decomposition of F

We consider this sum term by term for any two given level sequences S_1 and S_2. Let D_1 be the leveled degree sequence of $A \cup B$, and let D_2 be the leveled degree sequence of $(F_1 - A) \cup (F_2 - B)$. By applying the induction hypothesis with respect to $L(D)$ to A and B, we get

$$(n_{S_1^-}(A), n_{S_1^-}(B)) \blacktriangleleft (n_{S_1^-}(G_1(D_1)), n_{S_1^-}(G_2(D_1))) \tag{9.3}$$

and hence

$$n_{S_1^-}(G_1(D_1)) \geq \max\{n_{S_1^-}(A), n_{S_1^-}(B)\}. \tag{9.4}$$

On the other hand, applying the induction hypothesis with respect to r to $F_1 - A$ and $F_2 - B$ yields

$$(n_{S_2}(F_1 - A), n_{S_2}(F_2 - B)) \blacktriangleleft (n_{S_2}(G_1(D_2)), n_{S_2}(G_2(D_2))) \tag{9.5}$$

and consequently

$$n_{S_2}(G_1(D_2)) \geq \max\{n_{S_2}(F_1 - A), n_{S_2}(F_2 - B)\}. \tag{9.6}$$

Equations (9.4) and (9.6) imply

$$\begin{aligned} &n_{S_1^-}(G_1(D_1)) n_{S_2}(G_1(D_2)) \\ &\quad \geq (\max\{n_{S_1^-}(A), n_{S_1^-}(B)\}) \cdot (\max\{n_{S_2}(F_1 - A), n_{S_2}(F_2 - B)\}) \\ &\quad \geq \max\{n_{S_1^-}(A) n_{S_2}(F_1 - A), n_{S_1^-}(B) n_{S_2}(F_2 - B)\} \\ &\quad = \max\{n_{S_1}(A') n_{S_2}(F_1 - A), n_{S_1}(B') n_{S_2}(F_2 - B)\}. \end{aligned} \tag{9.7}$$

The relations (9.3) and (9.5) imply (see Lemma 7.10)

$$\begin{aligned} &n_{S_1^-}(G_1(D_1)) n_{S_2}(G_1(D_2)) + n_{S_1^-}(G_2(D_1)) n_{S_2}(G_2(D_2)) \\ &\qquad \geq n_{S_1^-}(A) n_{S_2}(F_1 - A) + n_{S_1^-}(B) n_{S_2}(F_2 - B). \end{aligned} \tag{9.8}$$

Let F^1 be the rooted forest whose first component F_1^1 is obtained by adding an edge joining the two roots of $G_1(D_2)$ and $G_1(D_1)$ and taking the root of $G_1(D_2)$ as root of F_1^1, and the second component F_2^1 is constructed analogously by adding an edge joining the roots

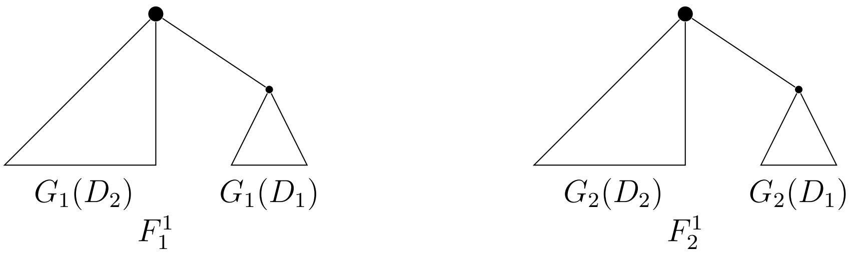


Figure 9.3: The rooted forest F^1

of $G_2(D_2)$ and $G_2(D_1)$ and taking the root of $G_2(D_2)$ as root of F_2^1 (if $G_2(D_1)$ is empty, then $F_2^1 = G_2(D_2)$). See Figure 9.3.

Since (9.7) and (9.8) are valid for arbitrary S_1 and S_2 satisfying the relation $S^- = S_1^- + S_2^-$, they imply that

$$n_S(F_1^1) \geq \max\{n_S(F_1), n_S(F_2)\}$$

and

$$n_S(F_1^1) + n_S(F_2^1) \geq n_S(F_1) + n_S(F_2).$$

We iterate this process to obtain a sequence of the form (with $F = F^0$)

$$(n_S(F_1^0), n_S(F_2^0)) \blacktriangleleft (n_S(F_1^1), n_S(F_2^1)) \blacktriangleleft \cdots \blacktriangleleft (n_S(F_1^K), n_S(F_2^K)).$$

The process terminates when it is no longer possible to find suitable branches A and B to replace. Clearly F^K must satisfy $F_1^K \triangleright F_2^K$. This means that F_1^K and F_2^K have the same leveled degree sequences as $G_1(D)$ and $G_2(D)$, respectively.

As our last step, we show that

$$n_S(F_1^K) \leq n_S(G_1(D)) \text{ and } n_S(F_2^K)) \leq n_S(G_2(D)), \tag{9.9}$$

which implies

$$(n_S(F_1^K), n_S(F_2^K)) \blacktriangleleft (n_S(G_1(D)), n_S(G_2(D))).$$

Recall that r_i is the degree of the root of F_i^K for $i = 1, 2$ and let $C(F_i^K) = \{H_1, \ldots, H_{r_i}\}$ and $C(G_i(D)) = \{H_1', \ldots, H_{r_i}'\}$. From the way F_i^K is formed we know that any $r_i - 1$ elements of $C(F_i^K)$ form a level greedy forest.

In particular, if $r_i \geq 3$, then any two elements of $C(F_i^K)$ form a level greedy forest. In view of Lemma 9.6, we conclude that the elements of $C(F_i^K)$ form a level greedy forest. Hence F_i^K is a level greedy tree, and (9.9) follows trivially.

If $r_i = 2$, we have:

- If $s_2 \geq 3$, then $n_S(F_i^K) = n_S(G_i(D)) = 0$.

- If $s_2 = 2$, let T and U be the trees obtained from F_i^K and $G_i(D)$ respectively by merging the root and its neighbors. Let $S' = (s_1, s_3, s_4, \ldots, s_n)$. Then we can use the outer induction hypothesis to get
$$n_S(F_i^K) = n_{S'}(T) \leq n_{S'}(U) = n_S(G_i(D)).$$
- If $s_2 = 1$, then we use again the (outer) induction hypothesis to get
$$n_S(F_i^K) = n_{S^-}(H_1) + n_{S^-}(H_2) \leq n_{S^-}(H_1') + n_{S^-}(H_2') = n_S(G_i(D)).$$
- The case $s_2 = 0$ is trivial.

If $r_i = 1$, the induction hypothesis of the first case gives us
$$n_S(F_i^K) = n_{S^-}(H_1) \leq n_{S^-}(H_1') = n_S(G_i(D))$$
if $s_2 \geq 1$, and
$$n_S(F_i^K) = n_S(G_i(D)) = 1$$
if $s_2 = 0$.

Finally, if $r_i = 0$, then the isolated vertex F_i^K is clearly a greedy tree. Thus we have shown (9.9) in all possible cases, so that
$$(n_S(F_1), n_S(F_2)) \blacktriangleleft (n_S(F_1^K), n_S(F_2^K)) \blacktriangleleft (n_S(G_1(D)), n_S(G_2(D))),$$
which completes the induction and thus our entire proof. □

A similar lemma also holds for edge-rooted trees in a completely analogous way.

Lemma 9.10 *Let D be the leveled degree sequence of an edge-rooted tree. For any $T \in \mathbb{T}_D$ we have $n_S(T) \leq n_S(G(D))$ for any level sequence $S = (s_1, s_2, \ldots, s_{L(D)})$.*

Proof. Let $D = ((i_{1,1}, i_{1,2}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$. If $s_1 \leq 1$, then edge roots do not play a role: using Lemma 9.8 we have
$$n_S(T) = n_S(T - \mathrm{r}(T)) \leq n_S(G(D) - \mathrm{r}(D)) = n_S(G(D)$$
because $T - \mathrm{r}(T)$ and $G(D) - \mathrm{r}(D)$ have the same leveled degree sequence $D' = ((i_{1,1}-1, i_{1,2}-1), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$ and $G(D) - \mathrm{r}(D) = G(D')$. The case $s_1 = 2$ is obtained by another application of Lemma 9.8, since in the case we have
$$n_S(G(D)) = n_{(s_1-1, s_2, \ldots, s_{L(D)})}(G(D''))$$
for $D'' = ((i_{1,1} + i_{1,2} - 2), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$, and
$$n_S(T) = n_{(s_1-1, s_2, \ldots, s_{L(D)})}(T'),$$

where T' is the tree obtained by merging the ends of the edge root to obtain a vertex root. Note that $G(D'')$ and T' are elements of $\mathbb{T}_{D''}$. Finally, if $s_1 \geq 3$, then clearly $n_S(T) = n_S(G(D)) = 0$. □

Our main result of this section follows as an immediate consequence of Lemmas 9.8 and 9.10 and Remark 7.6.

Theorem 9.11 *Among trees with a given degree sequence, $n_k(T)$ is maximized when T is the greedy tree.*

Proof. If it is possible to choose an edge or a vertex as root so that T is not level greedy, then we set T_1 as the level greedy tree with the same leveled degree sequence. We iterate this process, whenever it is possible, to obtain T_{l+1} from T_l just as T_1 is obtained from T. As we have seen in the proof of Theorem 8.10, no infinite loop is possible in this process. Hence there exists an integer m such that T_m is level greedy with respect to any choice of vertex or edge root. Such a T_m satisfies the "semi-regular" property defined in [78] (see Remark 7.6), and hence it is a greedy tree. By the two Lemmas 9.8 and 9.10, we know that

$$n_k(T) \leq n_k(T_1) \leq \cdots \leq n_k(T_m)$$

for any $k \geq 0$. □

Corollary 9.12 *Among forests with a given degree sequence, $n_k(F)$ is maximized when F is the greedy forest.*

Proof. Let F be a forest whose components are $F_1, \ldots, F_t$ ordered by non-increasing diameters. Whenever there is a possible choice of roots for F so that it has a leveled degree sequence B and it is not a level greedy forest, we have $n_k(F) \leq n_k(G(B))$. Hence we can choose F to be level greedy with respect to any choice of (vertex) roots.

Let v be a leaf end of a longest path in F_1. Assume that F_2 has a vertex w whose degree is larger than 1. Then the forest $F_1 \cup F_2$ considered to be rooted at v and w would not be level greedy: $\deg v < \deg w$ but the height of F_1 is larger than the height of F_2. Hence, F_1 is the only component of F that can possibly have vertices with degree greater than 1.

By Theorem 9.11 choosing F_1 to be greedy leaves unchanged or increases the number of subtrees of order k. With Remark 7.8, this completes the proof. □

9.3 Comparing different degree sequences

Comparing the greedy trees associated to different degree sequences, we will be able to determine the extremal trees with respect to the number of subtrees (of any given order) in a variety of different tree classes, cf. [94]. We use $D = (d_1, \ldots, d_n)$ and $B = (b_1, \ldots, b_n)$ to denote two different degree sequences (as opposed to leveled degree sequences unless otherwise mentioned) of trees.

The main goal of this section is to compare $n_k(G(B))$ and $n_k(G(D))$ if $B \preccurlyeq D$. It turns out that $n_k(G(B)) \leq n_k(G(D))$ in this case, from which a number of corollaries can be deduced.

For a vertex v in a rooted or edge-rooted tree T, let T_v denote the subtree induced by v and its descendants (Figure 9.4). Let $n_k(T, v)$ denote the number of subtrees of order k in T that contain the vertex v. The following lemma compares the values of n_k for all vertices on the same level of a level greedy tree.

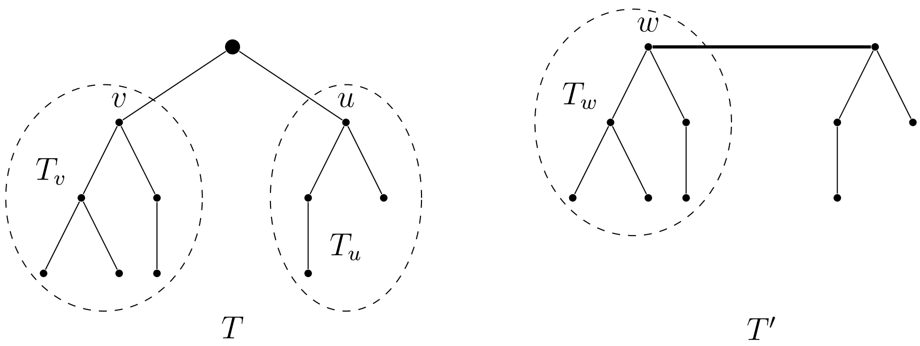


Figure 9.4: Definition of T_v

Lemma 9.13 *Let $D = ((i_{1,1}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$ be a leveled degree sequence of a tree. Then for all $1 \leq l \leq L(D)$ and $k \geq 1$ we have*

$$n_k(G(D), g_1^l) \geq n_k(G(D), g_2^l) \geq \cdots \geq n_k(G(D), g_{k_l-1}^l) \geq n_k(G(D), g_{k_l}^l).$$

Proof. Let $u = g_i^l$ and $v = g_j^l$ with $i < j$ be two vertices on the same level l of $T = G(D)$, and let w be their first (i.e., closest) common ancestor. We define a size-preserving injection from the set of all subtrees of $G(D)$ that contain v to the set of all subtrees of $G(D)$ that contain u. To this end, let u' and v' be the children of w for which $u \in T_{u'}$ and $v \in T_{v'}$. By the greedy construction, all vertices in $T_{u'}$ have greater or equal degree than all vertices in $T_{v'}$ on the same level. Hence there is an isomorphic embedding Φ of $T_{v'}$ into $T_{u'}$ that maps v to u, see Figure 9.5 for an example.

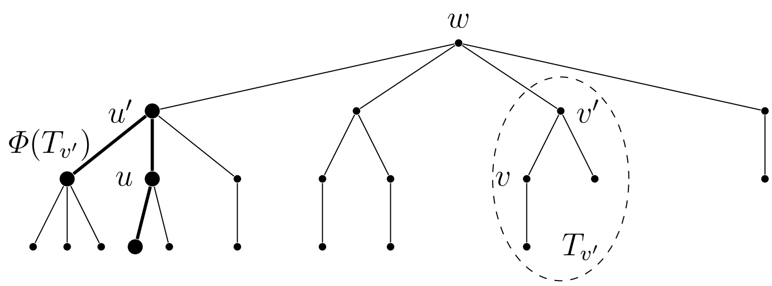


Figure 9.5: The tree $T_{v'}$ and its image $\Phi(T_{v'})$ (bold)

Let us now describe the size-preserving injection Ψ that maps a subtree R of $T = G(D)$ containing v to a subtree $\Psi(R)$ of $G(D)$ containing u. We distinguish three different cases:

1. If R contains both u and v, then we simply set $\Psi(R) = R$.

2. If R does not contain u and also does not contain w, then we set $\Psi(R) = \Phi(R)$.

3. If R does not contain u, but it does contain w, then let x be the first vertex (i.e., closest to w) on the path from w to u that is not contained in R, and let y be the vertex on the path from w to v that lies on the same level as x. Replace $R \cap T_y$ by $\Phi(R \cap T_y)$ (note that Φ maps the path from w to v to the path from w to u, thus y to x) to obtain $\Psi(R)$, see Figure 9.6.

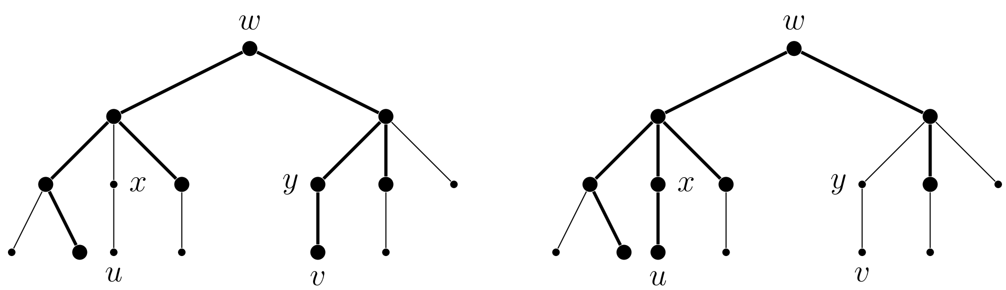


Figure 9.6: R (left) and R' (right) in Case (3)

Ψ inherits the properties of Φ of being injective and preserving the size of subtrees, so it follows immediately that $n_k(T, u) = n_k(G(D), g_i^l) \geq n_k(G(D), g_j^l) = n_k(T, v)$. $\square$

The same result holds for edge-rooted trees, and the proof is analogous:

Lemma 9.14 *Let* $D = ((i_{1,1}, i_{1,2}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$ *be a leveled degree sequence of an edge-rooted tree. Then we have*

$$n_k(G(D), g_1^l) \geq n_k(G(D), g_2^l) \geq \cdots \geq n_k(G(D), g_{k_l-1}^l) \geq n_k(G(D), g_{k_l}^l).$$

for any positive integer k *and* $1 \leq l \leq L(D)$*.*

We are now able to prove the main theorem of this section:

Theorem 9.15 *Let* $D = (d_1, \ldots, d_n)$ *and* $B = (b_1, \ldots, b_n)$ *be degree sequences of trees of the same order such that* $B \preccurlyeq D$*. Then for any positive integer* k *we have*

$$n_k(G(B)) \leq n_k(G(D)).$$

Proof. If $B = D$, the statement is trivial. Otherwise, there exists i_0 such that $d_{i_0} \neq b_{i_0}$. Since

$$\sum_{i=1}^{n} b_i = \sum_{i=1}^{n} d_i, \tag{9.10}$$

we know that the set $\{i : d_i \neq b_i\}$ must have at least two elements. Let $l = \min\{i : d_i \neq b_i\}$ and $m = \max\{i : d_i \neq b_i\}$. We must have $b_l < d_l$ and $b_m > d_m$. Note first that

$$B_1 = (b_1, \ldots, b_{l-1}, b_l + 1, b_{l+1}, \ldots, b_{m-1}, b_m - 1, b_{m+1}, \ldots, b_n)$$

is a valid degree sequence, because $b_{l-1} = d_{l-1} \geq d_l \geq b_l + 1$ and $b_{m+1} = d_{m+1} \leq d_m \leq b_m - 1$. It is easy to see that $B \preccurlyeq B_1$. Consider two vertices u and v in the greedy tree $G(B)$ such that $\deg u = b_l$ and $\deg v = b_m$.

Case 1: The length of the path in $G(B)$ joining u and v is even. Let w be the middle vertex of this path. Consider $G(B)$ as a level greedy tree whose root is w, then u and v are on the same level, say level h. We have $u = g_i^h$ and $v = g_j^h$ for some $i < j$. Without loss of generality, we may assume that j is the largest index such that $\deg g_j^h = b_m$ (otherwise, replace v by the vertex $g_{j'}^h$ which has this property). Let $x = g_r^{h+1}$ be a child of $v = g_j^h$, and let $H = G(B)_x$ be the branch rooted at x. Then $G(B) - H$ is still a level greedy tree (by maximality of j).

Consider the tree $T = G(B) - vx + ux$ with degree sequence B_1. Subtrees of $G(B)$ are still subtrees in T except for those that contain both v and x, but not u. On the other hand, we gain subtrees that contain u and x, but not v. This yields

$$n_k(T) - n_k(G(B)) = \sum_{k_1+k_2=k} n_{k_1}(H, x)\Big(n_{k_2}(G(B) - H, u) - n_{k_2}(G(B) - H, v)\Big),$$

which is nonnegative in view of Lemma 9.13. It follows that

$$n_k(G(B_1)) \geq n_k(T) \geq n_k(G(B))$$

for all $k > 0$.

Case 2: The length of the path in $G(B)$ joining u and v is odd. The argument is analogous to the previous case, but we use Lemma 9.14 instead of Lemma 9.13.

In either case, we have

$$n_k(G(B_1)) \geq n_k(G(B))$$

for all $k > 0$. We repeat this process to obtain a sequence of degree sequences $B_0 = B, B_1, B_2, \ldots, B_r = D$ such that

$$B = B_0 \preccurlyeq B_1 \preccurlyeq \cdots \preccurlyeq B_r = D$$

and

$$n_k(G(B)) = n_k(G(B_0)) \leq n_k(G(B_1)) \leq \cdots \leq n_k(G(B_r)) = n_k(G(D))$$

for all $k > 0$, which proves the theorem. □

A number of corollaries follow in the similar way as Corollaries 5.30 – 5.33 and 8.18 – 8.20 were obtained. Corollary 9.3, which has already been mentioned earlier, is such an instance. Let us mention a few more; the proofs are very similar, so we only give a proof of the first corollary.

Corollary 9.16 *For any tree T of order n, we have*

$$n_k(T) \leq \begin{cases} \binom{n-1}{k-1} & k > 1, \\ n & k = 1. \end{cases}$$

Proof. Note that the degree sequence $(n-1, 1, 1, \ldots, 1)$ of the star S_n majorizes all other degree sequences, and that

$$n_k(S_n) = \begin{cases} \binom{n-1}{k-1} & k > 1, \\ n & k = 1. \end{cases}$$

□

Corollary 9.17 *Among trees T of order n with s leaves, the number $n_k(T)$ is maximized by the greedy tree $G(s, 2, 2, \ldots, 2, 1, 1, \ldots, 1)$ (the number of 2s is $n - s - 1$, the number of 1s is s) for any $k \geq 1$.*

Corollary 9.18 *Among trees T of order n with independence number $\alpha \geq n/2$, the number $n_k(T)$ is maximized by the greedy tree*

$$G(\alpha, 2, 2, \dots, 2, 1, 1, \dots, 1)$$

(the number of 2s is $n - \alpha - 1$, the number of 1s is α) for any $k \geq 1$.

Corollary 9.19 *Among trees T of order n with matching number $\beta \leq n/2$, the number $n_k(T)$ is maximized by the greedy tree*

$$G(n - \beta, 2, 2, \dots, 2, 1, 1, \dots, 1)$$

(the number of 2s is $\beta - 1$, the number of 1s is $n - \beta$) for any $k \geq 1$.

9.4 Subtrees containing a given vertex

This section is devoted to subtrees containing a given vertex, which, as explained in Section 9.1, are strongly related to antichains in rooted trees. One of the consequences of Theorem 9.1 is the following:

Corollary 9.20 *Let $\rho_k(T)$ be the average number of subtrees of size k containing a randomly chosen vertex of T. The inequality*

$$\rho_k(T) \leq \rho_k(G(D))$$

holds for all $k \geq 1$ and all trees T of degree sequence D.

Proof. If we denote the order of T by n as usual, we have

$$\rho_k(T) = \frac{\sum_{v \in V(T)} n_k(T, v)}{n} = \frac{k n_k(T)}{n},$$

since each subtree of order k is counted k times. The desired inequality follows immediately. $\square$

It is natural to assume that the greedy tree also maximizes $n_k(T, v)$ if we choose v to be the canonical root. This fact, which has already been stated in the introduction (Theorem 9.4), is the main result of this section.

Proof of Theorem 9.4 Fix k, and let T be a rooted tree with degree sequence D. Consider a path $P = u_1 \dots u_m$ such that $m \geq 2$ and $u_1 = r(T)$ is the root of T. Let B be one of the branches attached to u_m by an edge such that $V(P) \cap V(B) = \varnothing$. Let T_i be the rooted tree obtained by removing B from u_m and attaching it to u_i for some $1 \leq i \leq m - 1$ (the root stays the same, see Figure 9.7).

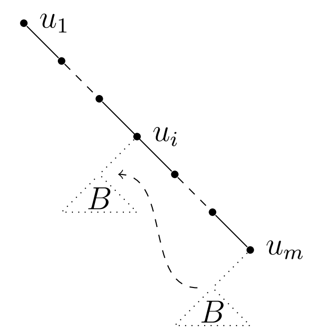


Figure 9.7: Moving the branch B from u_m to u_i

Then we have $n_k(T_i, r(T_i)) \geq n_k(T, r(T))$ for any $k \geq 1$ and $1 \leq i \leq m-1$: the number of subtrees which contain no vertex of B stays unchanged, and any subtree that contains the root as well as some vertices of B must contain the whole path P, thus it just gets transformed to a new subtree of the same order.

Now let T_M be a rooted tree with degree sequence D such that whenever a rooted tree T has degree sequence D, we always have $n_k(T_M, r(T_M)) \geq n_k(T, r(T))$. By the observation above, we can choose T_M such that $r(T_M)$ has maximum degree and the degrees of the vertices decrease as we move away from the root following a path. Hence if $D = (d_1, \ldots, d_n)$, then $\deg r(T_M) = d_1$ and there exists a neighbor v of $r(T_M)$ with $\deg v = d_2$. By Lemma 9.8, T_M can be chosen to be a level greedy tree. Let $n_k(T_M, e)$ be the number of subtrees of order k in T_M that contain the edge $e := r(T_M)v$. If T'_M is the component of $T_M - e$ that contains the root $r(T_M)$ (see Figure 9.8), then we have

$$n_k(T_M, r(T_M)) = n_k(T_M, e) + n_k(T'_M, r(T'_M)). \tag{9.11}$$

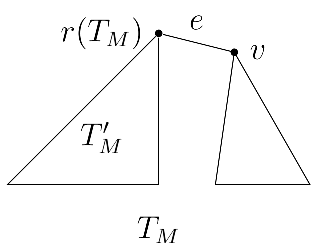


Figure 9.8: The tree T_M

By Lemma 9.10, we can reshuffle the branches in T_M to become a level greedy tree with edge root e, without decreasing $n_k(T_M, e)$ or $n_k(T'_M, r(T'_M))$: Note that the new T'_M obtained after reshuffling has the old one as a subgraph, given the fact that both of them are level greedy. Thus we also assume that T_M is level greedy with respect to the edge e.

For contradiction, assume that T_M (vertex rooted) is not isomorphic as rooted tree to $G(D)$. Then for some $i \geq 2$, there exist vertices u_i and u_{i+1} at levels i and $i+1$ respectively such that $\deg u_i < \deg u_{i+1}$. Using the fact that T_M is vertex rooted level greedy, we have the following where w_i and w_{i+1} are vertices at levels i and $i+1$ respectively:

Case 1: If $u_i, u_{i+1} \in V(T'_M)$, then there exists a vertex $w_{i+1} \in V(T_M - T'_M)$ such that $\deg u_i < \deg u_{i+1} \leq \deg w_{i+1}$.

Case 2: If $u_i, u_{i+1} \in V(T_M - T'_M)$, then there exists a vertex $w_i \in V(T'_M)$ such that $\deg w_i \leq \deg u_i < \deg u_{i+1}$. If level i of T'_M is already empty, we set $\deg w_i = 0$, and the argument that follows is still valid.

Case 3: If $u_i \in V(T_M - T'_M)$ and $u_{i+1} \in V(T'_M)$, then there exist vertices $w_i \in V(T'_M)$ and $w_{i+1} \in V(T_M - T'_M)$ such that

$$\deg w_i \leq \deg u_i < \deg u_{i+1} \leq \deg w_{i+1}.$$

The case that level i of T'_M is empty is treated in the same way as before. Hence all the three cases above can be reduced to the following fourth case:

Case 4: $u_i \in V(T'_M)$ and $u_{i+1} \in V(T_M - T'_M)$, but this contradicts the fact that T_M is level greedy as edge rooted tree with root e. $\square$

Before extending Theorem 9.4 a little further, we introduce the following related concepts.

Definition 9.21 A level greedy tree with leveled degree sequence

$$D = ((i_{1,1}), (i_{2,1}, \ldots, i_{2,k_2}), \ldots, (i_{n,1}, \ldots, i_{n,k_n}))$$

is called a *sliced greedy tree* if

$$\min\{i_{j,1}, \ldots, i_{j,k_j}\} \geq \max\{i_{j+1,1}, \ldots, i_{j+1,k_{j+1}}\}$$

for all $2 \leq j \leq n-1$. If furthermore we also have $i_{1,1} + 1 \geq \max\{i_{2,1}, \ldots, i_{2,k_2}\}$, then we say that the tree is a *branch greedy tree*.

In particular, a greedy tree is always a sliced greedy tree and a branch greedy tree. It is not hard to see that any sliced greedy tree can always be completed by adding further branches to turn it into a greedy tree, and that every branch of a greedy (sliced greedy, branch greedy) tree is branch greedy.

We now aim to extend Theorem 9.4, and show that among all rooted trees with given degree sequence and given degree of the root, the corresponding sliced greedy tree has the maximum number of subtrees of any given order containing the root.

Theorem 9.22 *Let $D = (d_1, \ldots, d_n)$ be a degree sequence of a tree. Let $\mathbb{T}_D^d$ be the set of all rooted trees with degree sequence D and root of degree d. Let $G(D,d)$ be the sliced greedy tree in $\mathbb{T}_D^d$. For any $T \in \mathbb{T}_D^d$ and for any positive integer k we have*

$$n_k(T, r(T)) \leq n_k(G(D,d), r(G(D,d))).$$

Proof. The case where $d = d_1$ coincides with Theorem 9.4. Hence, in the rest of the proof we assume that $d \leq d_2$. Since we know that

$$n_0(T, r(T)) = n_0(G(D,d), r(G(D,d))) = 0 \text{ and}$$
$$n_1(T, r(T)) = n_1(G(D,d), r(G(D,d))) = 1$$

we only have to check for the case where $k \geq 2$. We use an induction with respect to n. For the case where $n = 1, 2, 3$, the theorem clearly holds since $|\mathbb{T}_D^d| \leq 1$. Assume it holds whenever $1 \leq n \leq h$ for some integer $h \geq 3$. Now, consider the case where $n = h + 1$. By the same reasoning as in the first paragraph in the proof of Theorem 9.4, we can move branches in T closer to $r(T)$ without decreasing $n_k(T, r(T))$. Therefore, we can assume that if u is a neighbor of $r(T)$ and u' is any vertex of T which is in the branch of $r(T)$ containing u then $\deg u \geq \deg u'$. This and the induction hypothesis allow us to assume that each branch of $r(T)$ is branch greedy. In particular, $r(T)$ must have a neighbor v such that $\deg v = d_1$.

Let us start another induction on d. If $d = 1$, then for all $k \geq 1$ we have

$$\begin{aligned} n_k(T, r(T)) &= n_{k-1}(T - r(T), v) \\ &\leq n_{k-1}(G(D', d_1 - 1), r(G(D', d_1 - 1))) \\ &= n_k(G(D,d), r(G(D,d))), \end{aligned}$$

where D' is the degree sequence of $T - r(T)$.

Next we consider the case $d = 2$. Let T_1 and T_2 be the components of $T - r(T)$, where the neighbors of $r(T)$ are considered as roots and v is in $V(T_1)$. By Lemma 9.8, we can assume that $T_1 \rhd T_2$. If T and $G(D,d)$ have the same leveled degree sequence, then the two are isomorphic and we are done. Otherwise, there is an integer $i \geq 2$ such that there are two vertices u_i and u_{i+1} at the i^{th} and $(i+1)^{\text{th}}$ levels of T respectively, which satisfy

$$\deg u_i < \deg u_{i+1}. \tag{9.12}$$

We choose u_{i+1} such that its degree is maximum among all vertices at the $(i+1)^{\text{th}}$ level. Since both T_1 and T_2 are branch greedy, u_i and u_{i+1} must belong to different branches of $r(T)$. But it is impossible that

$u_i \in V(T_1)$ and $u_{i+1} \in V(T_2)$, since if we let $w_i \in V(T_2)$ be a vertex at the i^{th} level in T_2, then using the relation $T_1 \triangleright T_2$ and the fact that T_2 is branch greedy we would have $\deg u_i \geq \deg w_i \geq \deg u_{i+1}$, contradicting (9.12). Hence, we must have $u_i \in V(T_2)$ and $u_{i+1} \in V(T_1)$. Let T' be the tree obtained from T by merging v and $r(T)$ to become the new root, see Figure 9.9. Then we have (recall that we are assuming $k \geq 2$)

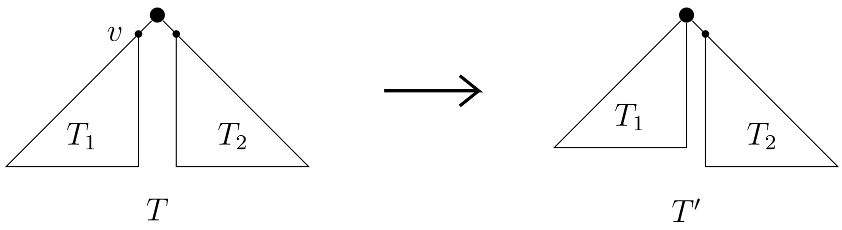


Figure 9.9: The tree T' in the proof of Theorem 9.22

$$n_k(T, r(T)) = n_{k-1}(T', r(T')) + n_{k-1}(T_2, r(T_2)). \tag{9.13}$$

Note that $n_{k-1}(T', r(T'))$ counts the subtrees of order k in T which contain the edge $vr(T)$, and $n_{k-1}(T_2, r(T_2))$ counts those that do not contain $vr(T)$.

Let $x_1, \ldots, x_{d_1-1}$ be the neighbors of $r(T_1)$ in T_1. We permute the vertices of T' to obtain a new, level greedy tree T'' with the same leveled degree sequence as T' but such that if $B_1, \ldots, B_{d_1}$ are the branches of $r(T'')$ containing $x_1, \ldots, x_{d_1-1}, r(T_2)$ respectively, then we have $B_{d_1} \triangleright \cdots \triangleright B_1$. Set $T_2'' := B_{d_1}$ and $T_1'' := T'' - B_{d_1}$. Let T''' be a tree obtained from T by replacing T_1 and T_2 by T_1'' and T_2'', respectively. Note that T_2 is isomorphic to a subgraph of T_2''. Let T^1 be the level greedy tree with the same leveled degree sequence as T'''. From Lemma 9.8, we deduce that $n_{k-1}(T', r(T')) \leq n_{k-1}(T'', r(T''))$, $n_{k-1}(T_2, r(T_2)) \leq n_{k-1}(T_2'', r(T_2''))$ and consequently

$$\begin{aligned} n_k(T, r(T)) &= n_{k-1}(T', r(T')) + n_{k-1}(T_2, r(T_2)) \\ &\leq n_{k-1}(T'', r(T'')) + n_{k-1}(T_2'', r(T_2'')) \\ &= n_k(T''', r(T''')) \leq n_k(T^1, r(T^1)). \end{aligned}$$

Along this process, at least one vertex with maximum degree at the $(i+1)^{\text{th}}$ level in T (hence the same degree as u_{i+1}) is transfered to the i^{th} level in T'', T''' and T^1. We can iterate the same process until we reach a tree that is isomorphic to $G(D, d)$.

Now we can resume our induction with respect to the degree d. Assume that the Theorem holds for $d = m$ for some integer $m \geq 2$. Now, consider the case where $d = m+1$. By the induction hypothesis with respect to d, we can assume that whenever B is a branch of $r(T)$ then $T - B$ is a sliced greedy tree. For any vertex u_i at the i^{th} level and u_{i+1} at the $(i+1)^{\text{th}}$ level in T for some $i \geq 2$, there exists a branch B of $r(T)$ such that $\{u_i, u_{i+1}\} \cap V(B) = \varnothing$, since $d \geq 3$. Since $T - B$ is a sliced greedy tree, we must have $\deg u_i \geq \deg u_{i+1}$. This means that T has the same leveled degree sequence as $G(D, d)$. Hence, by Lemma 9.8 we get $n_k(G(D,d), r(G(D,d))) \geq n_k(T, r(T))$. □

Recall that subtrees containing the root correspond to antichains[1] when a rooted tree is regarded as a Hasse diagram[1] of a poset. Since the cardinality of the antichain corresponds to the number of leaves (counting the root as a leaf only if it is the only vertex of the subtree), it is natural to ask whether similar statements as Theorem 9.4 and Theorem 9.22 remain true if the number of subtrees with a fixed number l of leaves is considered instead. This turns out to be the case.

For any (vertex) rooted forest F, let $\eta_l(F)$ denote the number of subtrees in F which contain one of the roots and have l leaves (as before, the root is only counted as a leaf if it is the only vertex). It is convenient to set $\eta_0(F) = 1$ and $\eta_l(F) = 0$ for negative l. Moreover, it is easy to see that $\eta_1(F) = |F|$ only depends on the order of F, so we will focus on the case $l \geq 2$ in the following.

Lemma 9.23 *Let D be a given leveled degree sequence of a two-component forest. For any positive integer l and for any $F \in \mathbb{F}_D$ we have*

$$(\eta_l(F_1), \eta_l(F_2)) \blacktriangleleft (\eta_l(G_1(D)), \eta_l(G_2(D))). \tag{9.14}$$

Similarly, if D is a leveled degree sequence of a (vertex) rooted tree, then for any $T \in \mathbb{T}_D$ and for any positive integer l we have

$$\eta_l(T) \leq \eta_l(G(D)). \tag{9.15}$$

Proof. By the same reasoning as in Remark 9.9, we only have to show (9.14), where we allow F_2 to be empty. We use an induction on $L(D)$. The cases $L(D) = 1, 2$ are trivial, since the degree sequence uniquely characterizes the tree in these cases. Assume that the lemma is true whenever $L(D) \leq k$ for some integer $k \geq 2$, and suppose that $L(D) = k + 1$. Let r_i be the degree of the root of F_i for $i = 1, 2$, and assume that $r_1 \geq r_2$. We start a new induction, this time with respect to r_1. For $r_1 = 1$ and any $l \geq 2$, we have

$$\eta_l(G_1(D)) = \eta_l(G_1(D) - r(G_1(D))) \text{ and } \eta_l(F_1) = \eta_l(F_1 - r(F_1)),$$

[1] See Appendix C

for non-empty $G_2(D)$ we get

$$\eta_l(G_2(D)) = \eta_l(G_2(D) - r(G_2(D))) \text{ and } \eta_l(F_2) = \eta_l(F_2 - r(F_2)),$$

and for empty $G_2(D)$ we still have

$$\eta_l(F_2) = \eta_l(F_2 - r(F_2)) = 0 \text{ and } \eta_l(G_2(D)) = \eta_l(G_2(D) - r(G_2(D))) = 0.$$

Hence (9.14) follows (by the induction hypothesis) from

$$(\eta_l(F_1 - r(F_1)), \eta_l(F_2 - r(F_2))) \blacktriangleleft (\eta_l(G_1(D) - r(G_1(D))), \eta_l(G_2(D) - r(G_2(D)))).$$

Assume (9.14) holds whenever $r_1 \leq m$ for some $m \geq 1$, and let $r_1 = m + 1$. Let A and B be subtrees of F_1 and F_2 as shown in Figure 9.2, where B is empty if F_2 is an isolated vertex or if F_2 is empty. Then the following relation holds:

$$\begin{aligned}\eta_l(F) = {} & \eta_{l-1}(A)(\eta_1(F_1 - A) - 1) + \eta_{l-1}(B)(\eta_1(F_2 - B) - 1) \\ & + \sum_{\substack{l_1, l_2 \geq 0, l_2 \neq 1 \\ l_1 + l_2 = l}} \eta_{l_1}(A)\eta_{l_2}(F_1 - A) + \eta_{l_1}(B)\eta_{l_2}(F_2 - B).\end{aligned}$$

This follows from the fact that the l leaves have to be divided into l_1 leaves in A (or B) and l_2 leaves in $F_1 - A$ (or $F_2 - B$) respectively. The only exception is the case $l_2 = 1$: the subtree of $F_1 - A$ ($F_2 - B$) that only consists of the root counts with 0 leaves. The next step of the proof follows the same lines as the proof of Lemma 9.8, so we skip the details: the induction hypothesis shows that each term in the above sum is maximized when A, B, $F_1 - A$ and $F_2 - B$ are components of greedy forests. This argument gives us a forest F^K whose components F_1^K and F_2^K have the same leveled degree sequences as $G_1(D)$ and $G_2(D)$ respectively, and which satisfies

$$(\eta_l(F_1), \eta_l(F_2)) \blacktriangleleft (\eta_l(F_1^K), \eta_l(F_2^K)).$$

Moreover, if d_i denotes the root degree of F_i^K, we can assume that any $d_i - 1$ elements of $C(F_i^K)$ form a level greedy forest. If $d_i \neq 2$, then it follows immediately that F_i^K is a greedy tree (as in Lemma 9.8). Thus we only consider the case $d_i = 2$.

Let $C(F_i^K) = \{H_1, H_2\}$ and $C(G_i(D)) = \{H_1', H_2'\}$, and let $F_i'^K$ and $G_i'(D)$ be, respectively, the trees obtained from F_i^K by merging $r(F_i^K)$ with its two neighbors and from $G_i(D)$ by merging $r(G_i(D))$ with its neighbors. Then the induction hypothesis with respect to $L(D)$ applied to $H_1 \cup H_2$ yields

$$\begin{aligned}\eta_2(F_i^K) &= \eta_2(F_i'^K) + \eta_1(H_1) + \eta_1(H_2) - 1 \\ &\leq \eta_2 G_i'(D)) + \eta_1(H_1') + \eta_1(H_2') - 1 = \eta_2(G_i(D)),\end{aligned}$$

where the term -1 is due to an over-count of the subtree having the two neighbors of $r(F_i)$ or of $r(G_i(D))$ as leaves. For all $l \geq 3$ we have

$$\begin{aligned} \eta_l(F_i^K) &= \eta_l(F_i'^K) + \eta_{l-1}(H_1) + \eta_{l-1}(H_2) \\ &\leq \eta_l(G_i'(D)) + \eta_{l-1}(H_1') + \eta_{l-1}(H_2') = \eta_l(G_i(D)). \end{aligned} \tag{9.16}$$

Note here that $\eta_l(F_i'^K)$ is the number of subtrees in F_i^K containing $r(F_i^K)$ and having l leaves none of which is a neighbor of $r(F_i^K)$, while $\eta_{l-1}(H_1)$ and $\eta_{l-1}(H_2)$ count subtrees with l leaves one of which is a neighbor of $r(F_i^K)$. □

It turns out that, as one would expect intuitively, moving a branch to be closer to the root does not decrease η_l of a tree for any l.

Lemma 9.24 *Let T be a vertex rooted tree, and let $P = u_1 \dots u_m$ ($m \geq 2$) be a path starting at the root (i.e., $u_1 = r(T)$). Let B be a branch attached by an edge to u_m such that $V(B) \cap V(P) = \varnothing$. Let v be the neighbor of u_m in B, and let T' be a tree obtained from T by removing the edge vu_m and adding a new edge vu_{m-1} (see Figure 9.7 with $i = m-1$). Then we have $\eta_l(T) \leq \eta_l(T')$.*

Proof. Clearly we have $\eta_l(T - B) = \eta_l(T' - B)$. Hence we only have to compare the number of subtrees containing some vertices of B. To any subtree S in T which contains $r(T)$ and such that $V(S) \cap V(B) \neq \varnothing$, we associate $f(S)$ defined as follows:

- $f(S) = S - vu_m + vu_{m-1}$ if u_m is not a leaf in $S - vu_m + vu_{m-1}$;
- otherwise we have $f(S) = S - u_m + vu_{m-1}$.

Clearly $f(S)$ is a subtree of T', it has the same number of leaves as S, and if $f(S) = f(S')$ then we also have $S = S'$: add u_m if it is missing, add vu_m and remove vu_{m-1} to obtain S and S' from $F(S)$ and $F(S')$, respectively. Hence f is injective, and we are done. □

Theorem 9.25 *For all rooted trees T with degree sequence D and for all positive integers l we have $\eta_l(G(D)) \geq \eta_l(T)$, where $G(D)$ is rooted in the canonical way as in Definition 7.5.*

Proof. Let T_M be a rooted tree with degree sequence $D = (d_1, \dots, d_n)$, and such that for all rooted trees T with degree sequence D we have $\eta_l(T_M) \geq \eta_l(T)$. By the two Lemmas 9.23 and 9.24, we can choose T_M to be level greedy with $\deg r(T_M) = d_1$, and such that the degrees decrease along each path starting from $r(T_M)$. Hence $r(T_M)$ has a neighbor v whose degree is d_2. Let T_M' be the component of $T_M - vr(T_M)$ that contains $r(T_M)$, and let T_M'' be the tree obtained from T_M by merging v and $r(T_M)$ to become the new root. Let A be the set of subtrees of T_M containing $r(T_M)$ and having l leaves such that one of them is v, and let B be the set of the subtrees of T_M containing $r(T_M)$ and l leaves none of which is v. Then we have

$$\eta_l(T_M) = |A| + |B| = \eta_{l-1}(T'_M) + \eta_l(T''_M)$$

for $l \geq 3$; for $l = 2$ we have $|A| = \eta_1(T'_M) - 1$ (the -1 is for the subtree that only consists of $r(T_M)$ and v) and hence

$$\eta_2(T_M) = |A| + |B| = \eta_1(T'_M) + \eta_2(T''_M) - 1.$$

If we reshuffle the branches of v and $r(T_M)$ such that T_M considered as edge rooted tree with $vr(T_M)$ as root is level greedy, the values of $\eta_{l-1}(T'_M)$ and $\eta_l(T''_M)$ will not decrease. This is because T''_M would then be level greedy as well, and the new T'_M has the old one as subgraph (the fact that T'_M is level greedy is used here). By exactly the same reason as in the last paragraph in the proof of Theorem 9.4, we can show that it is impossible to find in T_M vertices u_i and u_{i+1} at levels i and $i+1$ respectively such that $\deg u_i < \deg u_{i+1}$. This implies that $T_M = G(D)$. □

More generally, we also have a theorem analogous to Theorem 9.22:

Theorem 9.26 *Let T be a rooted tree with degree sequence $D = (d_1, \ldots, d_n)$ and root degree $\deg r(T) = d$. Then for any positive integer l, we have $\eta_l(T) \leq \eta_l(G(D, d))$, where $G(D, d)$ denotes the sliced greedy tree whose root has degree d and whose degree sequence is D.*

Proof. See Theorem 9.25 for the case where $d = d_1$. As mentioned earlier, the case $l = 1$ is trivial, so we assume that $d \leq d_2$ and $l \geq 2$. By Lemma 9.24, we can restrict ourselves to the case where $r(T)$ has a neighbor v with $\deg v = d_1$. Let T' be the tree obtained from T by merging $r(T)$ and v and let T_2 be the connected component of $T - vr(T)$ that contains $r(T)$. Then for all $l \geq 3$ we have

$$\eta_l(T) = \eta_l(T') + \eta_{l-1}(T_2) \tag{9.17}$$

as in the previous proof (for $l = 2$, we have to subtract 1). Here, $\eta_l(T')$ counts the number of subtrees of T containing $r(T)$ and having l leaves none of which is v, and $\eta_{l-1}(T_2)$ counts the subtrees of T containing $r(T)$ such that one of its l leaves is v. The rest of the proof is exactly as we have seen in the proof of Theorem 9.22, but we use (9.17) instead of (9.13), and we use Lemma 9.23 in the place of Lemma 9.8. □

Appendices

Appendix A

Approximate values of energy of tripods

Let us denote by $\mathfrak{d}(i, j)$ the limit

$$\lim_{n\to\infty} \mathrm{En}(T(i, j, n-i-j-1)) - \mathrm{En}(H(2, 2, 2, 2, n))$$

for all integers $i \leq j$. Then we have the following values rounded to six decimal places:

$\mathfrak{d}(2,2) \approx 0.065993$	$\mathfrak{d}(2,7) \approx 0.040329$	$\mathfrak{d}(4,4) \approx 0.037612$
$\mathfrak{d}(2,5) \approx 0.037493$	$\mathfrak{d}(4,6) \approx 0.031345$	$\mathfrak{d}(2,3) \approx 0.030513$
$\mathfrak{d}(4,8) \approx 0.028282$	$\mathfrak{d}(4,18) \approx 0.023946$	$\mathfrak{d}(6,6) \approx 0.023909$
$\mathfrak{d}(4,20) \approx 0.023683$	$\mathfrak{d}(4,15) \approx 0.020244$	$\mathfrak{d}(6,8) \approx 0.020135$
$\mathfrak{d}(4,13) \approx 0.019647$	$\mathfrak{d}(4,11) \approx 0.018763$	$\mathfrak{d}(6,10) \approx 0.017948$
$\mathfrak{d}(4,9) \approx 0.017378$	$\mathfrak{d}(6,12) \approx 0.016566$	$\mathfrak{d}(8,8) \approx 0.015891$
$\mathfrak{d}(6,14) \approx 0.015637$	$\mathfrak{d}(4,7) \approx 0.015026$	$\mathfrak{d}(6,16) \approx 0.014983$
$\mathfrak{d}(6,18) \approx 0.014504$	$\mathfrak{d}(6,26) \approx 0.013472$	$\mathfrak{d}(8,10) \approx 0.013381$
$\mathfrak{d}(6,28) \approx 0.013329$	$\mathfrak{d}(6,39) \approx 0.011775$	$\mathfrak{d}(8,12) \approx 0.011767$
$\mathfrak{d}(6,37) \approx 0.011718$	$\mathfrak{d}(6,23) \approx 0.010881$	$\mathfrak{d}(8,14) \approx 0.010667$
$\mathfrak{d}(10,10) \approx 0.010636$	$\mathfrak{d}(6,21) \approx 0.010629$	$\mathfrak{d}(4,5) \approx 0.010537$
$\mathfrak{d}(6,19) \approx 0.010306$	$\mathfrak{d}(8,16) \approx 0.009883$	$\mathfrak{d}(6,17) \approx 0.009881$
$\mathfrak{d}(6,15) \approx 0.009306$	$\mathfrak{d}(8,18) \approx 0.009304$	$\mathfrak{d}(8,20) \approx 0.008864$
$\mathfrak{d}(10,12) \approx 0.008849$	$\mathfrak{d}(8,22) \approx 0.008523$	$\mathfrak{d}(6,13) \approx 0.008504$
$\mathfrak{d}(8,24) \approx 0.008252$	$\mathfrak{d}(8,30) \approx 0.007708$	$\mathfrak{d}(10,14) \approx 0.007616$
$\mathfrak{d}(8,32) \approx 0.007584$	$\mathfrak{d}(8,34) \approx 0.00748$	$\mathfrak{d}(8,36) \approx 0.007391$
$\mathfrak{d}(6,11) \approx 0.007334$	$\mathfrak{d}(8,38) \approx 0.007314$	$\mathfrak{d}(8,56) \approx 0.006934$
$\mathfrak{d}(12,12) \approx 0.006928$	$\mathfrak{d}(8,58) \approx 0.006911$	$\mathfrak{d}(8,86) \approx 0.006732$
$\mathfrak{d}(10,16) \approx 0.006729$	$\mathfrak{d}(8,88) \approx 0.006725$	$\mathfrak{d}(8,49) \approx 0.006104$
$\mathfrak{d}(10,18) \approx 0.006069$	$\mathfrak{d}(8,47) \approx 0.006066$	$\mathfrak{d}(8,33) \approx 0.005602$
$\mathfrak{d}(12,14) \approx 0.005592$	$\mathfrak{d}(10,20) \approx 0.005564$	$\mathfrak{d}(6,9) \approx 0.005533$

$\mathfrak{d}(8,31) \approx 0.005486$	$\mathfrak{d}(8,29) \approx 0.005347$	$\mathfrak{d}(8,27) \approx 0.005181$
$\mathfrak{d}(10,22) \approx 0.005169$	$\mathfrak{d}(8,25) \approx 0.004979$	$\mathfrak{d}(10,24) \approx 0.004855$
$\mathfrak{d}(8,23) \approx 0.00473$	$\mathfrak{d}(12,16) \approx 0.004622$	$\mathfrak{d}(10,26) \approx 0.0046$
$\mathfrak{d}(1,2) \approx 0.004585$	$\mathfrak{d}(8,21) \approx 0.004418$	$\mathfrak{d}(10,28) \approx 0.004391$
$\mathfrak{d}(10,30) \approx 0.004217$	$\mathfrak{d}(14,14) \approx 0.004172$	$\mathfrak{d}(10,32) \approx 0.004071$
$\mathfrak{d}(8,19) \approx 0.004018$	$\mathfrak{d}(10,34) \approx 0.003947$	$\mathfrak{d}(12,18) \approx 0.003895$
$\mathfrak{d}(10,36) \approx 0.003841$	$\mathfrak{d}(10,44) \approx 0.00354$	$\mathfrak{d}(8,17) \approx 0.003497$
$\mathfrak{d}(10,46) \approx 0.003486$	$\mathfrak{d}(10,52) \approx 0.003358$	$\mathfrak{d}(12,20) \approx 0.003336$
$\mathfrak{d}(10,54) \approx 0.003324$	$\mathfrak{d}(10,70) \approx 0.003141$	$\mathfrak{d}(14,16) \approx 0.003135$
$\mathfrak{d}(10,72) \approx 0.003126$	$\mathfrak{d}(10,182) \approx 0.002896$	$\mathfrak{d}(12,22) \approx 0.002896$
$\mathfrak{d}(10,184) \approx 0.002895$	$\mathfrak{d}(10,175) \approx 0.0028$	$\mathfrak{d}(8,15) \approx 0.002799$
$\mathfrak{d}(10,173) \approx 0.002799$	$\mathfrak{d}(10,69) \approx 0.002548$	$\mathfrak{d}(6,7) \approx 0.002545$
$\mathfrak{d}(12,24) \approx 0.002544$	$\mathfrak{d}(10,67) \approx 0.002531$	$\mathfrak{d}(10,53) \approx 0.002355$
$\mathfrak{d}(14,18) \approx 0.002354$	$\mathfrak{d}(10,51) \approx 0.002318$	$\mathfrak{d}(10,49) \approx 0.002277$
$\mathfrak{d}(12,26) \approx 0.002258$	$\mathfrak{d}(10,47) \approx 0.002231$	$\mathfrak{d}(10,41) \approx 0.002055$
$\mathfrak{d}(16,16) \approx 0.002043$	$\mathfrak{d}(12,28) \approx 0.002022$	$\mathfrak{d}(10,39) \approx 0.001979$
$\mathfrak{d}(10,37) \approx 0.001892$	$\mathfrak{d}(8,13) \approx 0.001834$	$\mathfrak{d}(12,30) \approx 0.001825$
$\mathfrak{d}(10,35) \approx 0.001792$	$\mathfrak{d}(14,20) \approx 0.001749$	$\mathfrak{d}(10,33) \approx 0.001674$
$\mathfrak{d}(12,32) \approx 0.001659$	$\mathfrak{d}(10,31) \approx 0.001536$	$\mathfrak{d}(12,34) \approx 0.001518$
$\mathfrak{d}(12,36) \approx 0.001397$	$\mathfrak{d}(10,29) \approx 0.001373$	$\mathfrak{d}(12,38) \approx 0.001292$
$\mathfrak{d}(14,22) \approx 0.001271$	$\mathfrak{d}(16,18) \approx 0.001216$	$\mathfrak{d}(12,40) \approx 0.001201$
$\mathfrak{d}(10,27) \approx 0.001176$	$\mathfrak{d}(12,42) \approx 0.001122$	$\mathfrak{d}(12,44) \approx 0.001052$
$\mathfrak{d}(12,46) \approx 0.00099$	$\mathfrak{d}(10,25) \approx 0.000938$	$\mathfrak{d}(12,48) \approx 0.000934$
$\mathfrak{d}(14,24) \approx 0.000887$	$\mathfrak{d}(12,50) \approx 0.000885$	$\mathfrak{d}(12,64) \approx 0.000651$
$\mathfrak{d}(10,23) \approx 0.000646$	$\mathfrak{d}(12,66) \approx 0.000629$	$\mathfrak{d}(12,70) \approx 0.000589$
$\mathfrak{d}(14,26) \approx 0.000573$	$\mathfrak{d}(16,20) \approx 0.000573$	$\mathfrak{d}(12,72) \approx 0.000571$
$\mathfrak{d}(12,92) \approx 0.00045$	$\mathfrak{d}(8,11) \approx 0.000443$	$\mathfrak{d}(12,94) \approx 0.000442$
$\mathfrak{d}(12,130) \approx 0.000351$	$\mathfrak{d}(18,18) \approx 0.00035$	$\mathfrak{d}(12,132) \approx 0.000348$
$\mathfrak{d}(12,162) \approx 0.000314$	$\mathfrak{d}(14,28) \approx 0.000314$	$\mathfrak{d}(12,164) \approx 0.000312$
$\mathfrak{d}(12,224) \approx 0.000282$	$\mathfrak{d}(10,21) \approx 0.000281$	$\mathfrak{d}(12,226) \approx 0.000281$
$\mathfrak{d}(12,219) \approx 0.000207$	$\mathfrak{d}(3,4) \approx 0.000206$	$\mathfrak{d}(12,217) \approx 0.000206$
$\mathfrak{d}(12,111) \approx 0.000101$	$\mathfrak{d}(14,30) \approx 0.000097$	$\mathfrak{d}(12,109) \approx 0.000096$
$\mathfrak{d}(12,99) \approx 0.000065$	$\mathfrak{d}(16,22) \approx 0.000063$	$\mathfrak{d}(12,97) \approx 0.000058$
$\mathfrak{d}(12,85) \approx 0.000005$		

Appendix B

The values of $\Lambda(n)$ and $\Omega(n)$ for n ranging from 1 to 100.

Even n						Odd n					
n	$\Lambda(n)$	$\Omega(n)$	n	$\Lambda(n)$	$\Omega(n)$	n	$\Lambda(n)$	$\Omega(n)$	n	$\Lambda(n)$	$\Omega(n)$
10	5	5	56	107	24	11	5	4	57	106	40
12	7	4	58	113	25	13	7	6	59	113	41
14	9	3	60	119	26	15	8	7	61	121	43
16	12	4	62	126	27	17	10	4	63	128	46
18	14	5	64	132	28	19	12	5	65	135	49
20	17	6	66	138	29	21	15	5	67	143	52
22	21	7	68	142	30	23	20	6	69	148	55
24	25	8	70	149	31	25	23	6	71	154	57
26	28	9	72	155	32	27	26	9	73	158	59
28	33	10	74	161	33	29	29	10	75	163	61
30	37	11	76	167	34	31	33	15	77	169	64
32	42	12	78	172	35	33	36	16	79	173	66
34	48	13	80	179	36	35	42	24	81	177	68
36	51	14	82	186	37	37	47	26	83	183	70
38	56	15	84	191	38	39	52	26	85	188	72
40	61	16	86	197	39	41	57	28	87	194	74
42	68	17	88	202	40	43	62	29	89	198	76
44	72	18	90	207	41	45	67	29	91	206	78
46	78	19	92	213	42	47	72	30	93	213	81
48	85	20	94	219	43	49	78	31	95	220	83
50	90	21	96	225	112	51	83	32	97	226	85
52	97	22	98	231	115	53	90	33	99	230	87
54	102	23	100	236	118	55	99	34			

Appendix C

Hasse diagrams of posets

A partial order is a binary relation, say "$\leq$", over a set X which satisfies the following conditions for all x, y and z in X:

- $x \leq x$ (reflexivity);
- if $x \leq y$ and $y \leq x$ then $x = y$ (antisymmetry);
- if $x \leq y$ and $y \leq z$ then $x \leq z$ (transitivity).

A partially ordered set (also called a poset) is a set with a partial order.

We say that y *covers* x in a poset $(X, \leq)$ if $y, x \in X$, $y \neq x$ and there exists no $z \in X - \{y, x\}$ such that $x \leq z \leq y$. The *Hasse diagram* of a finite poset X is the graph whose vertices are the elements of X, and $\{x, y\}$ is an edge if and only if $x \neq y$ and x covers y or y covers x. If y covers x and $x \neq y$ then y appears "above" x. See Figure C.1, where X is the power set of $\{a, b, c\}$, and the partial order is the inclusion.

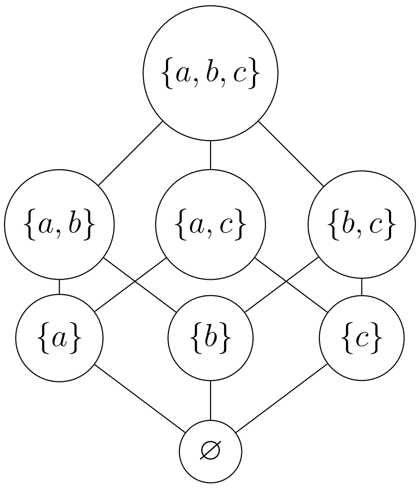


Figure C.1: Hasse diagram of the power set of $\{a, b, c\}$ ordered by inclusion

Two elements x and y of X are *incomparable* in a poset $(X, \leq)$ if $x \nleq y$ and $y \nleq x$. An *antichain* is a set of pairwise incomparable elements. For example in the poset represented in Figure C.1, $\{\{a\}, \{b\}, \{c\}\}$ is an antichain, and so is $\{\{a\}, \{b, c\}\}$.

Bibliography

[1] E. O. D. Andriantiana, The number of independent subsets and the energy of trees, Master thesis at Stellenbosch University (2010).

[2] E. O. D. Andriantiana, Unicyclic bipartite graphs with maximum energy, MATCH Commun. Math. Comput. Chem. 66 (3) (2011) 913–926.

[3] E. O. D. Andriantiana, More trees with large energy, MATCH Commun. Math. Comput. Chem. 68 (2012) 675–695.

[4] E. O. D. Andriantiana, Energy, Hosoya index and Merrifield–Simmons index of trees with prescribed degree sequence, Discrete Appl. Math. 161 (2013) 724–741.

[5] E. O. D. Andriantiana, S. Wagner, On the number of independent subsets in trees with restricted degrees, Math. Comput. Model. 53 (2011) 678–683.

[6] E. O. D. Andriantiana, S. Wagner, Unicyclic graphs with large energy, Linear Algebra and its Applications 435 (6) (2011) 1399–1414.

[7] E. O. D. Andriantiana, S. Wagner, Spectral moment of trees with given degree sequence, submitted (2013).

[8] E. O. D. Andriantiana, S. Wagner, H. Wang, Greedy trees, subtrees and antichains, Electron. J. Combin. 20 (3) (2013) 28 pp.

[9] T. Bıyıkoğlu, M. Hellmuth, J. Leydold, Largest eigenvalues of the discrete p-Laplacian of trees with degree sequences, Electron. J. Linear Algebra 18 (2009) 202–210.

[10] T. Bıyıkoğlu, J. Leydold, Graphs with given degree sequence and maximal spectral radius, Electron. J. Combin. 15 (1) (2008) Research Paper 119, 9 pp.

[11] G. Caporossi, D. Cvetković, I. Gutman, P. Hansen, Variable neighborhood search for extremal graphs. 2. finding graphs with extremal energy, J. Chem. Inf. Comput. Sci. 39 (1999) 984–996.

[12] A. Chen, A. Chang, W. C. Shiu, Energy ordering of unicycle graphs, MATCH Commun. Math. Comput. Chem. 55 (2006) 95–102.

[13] C. Coulson, B. O'Leary, R. Mallion, Hückel theory for organic chemists, Academic Press, 1978.

[14] C. A. Coulson, On the calculation of the energy in unsaturated hydrocarbon molecules, Mathematical Proceedings of the Cambridge Philosophical Society 36 (1940) 201–203.

[15] D. Cvetković, M. Doob, H. Sachs, Spectra of graphs – Theory and application, 3rd ed., Johann Ambrosius Barth Verlag, Heidelberg, Leipzig, 1995.

[16] R. Diestel, Graph theory, vol. 173 of Graduate Texts in Mathematics, 3rd ed., Springer-Verlag, Berlin, 2005.

[17] Z. Du, B. Zhou, The Estrada index of trees, Linear Algebra Appl. 435 (10) (2011) 2462–2467.

[18] Z. Du, B. Zhou, On the Estrada index of graphs with given number of cut edges, Electron. J. Linear Algebra 22 (2011) 586–592.

[19] Z. Du, B. Zhou, The Estrada index of unicyclic graphs, Linear Algebra Appl. 436 (9) (2012) 3149–3159.

[20] Z. Du, B. Zhou, R. Xing, On maximum Estrada indices of graphs with given parameters, Linear Algebra Appl. 436 (9) (2012) 3767–3772.

[21] E. Estrada, Characterization of 3d molecular structure, Chemical Physics Letters 319 (5–6) (2000) 713–718.

[22] E. Estrada, Characterization of the folding degree of proteins, Bioinformatics 18 (5) (2002) 697–704.

[23] E. Estrada, Characterization of the amino acid contribution to the folding degree of proteins, Proteins 54 (2004) 727–737.

[24] E. Estrada, Topological structural classes of complex networks, Phys. Rev. E 75 (2007) 016103.

[25] E. Estrada, J. A. Rodríguez-Velázquez, M. Randić, Atomic branching in molecules, International Journal of Quantum Chemistry 106 (4) (2006) 823–832.

[26] M. Fischermann, A. Hoffmann, D. Rautenbach, L. Székely, L. Volkmann, Wiener index versus maximum degree in trees, Discrete Appl. Math. 122 (2002) 127–137.

[27] Y. D. Gao, H. Hosoya, Topological index and thermodynamic properties. IV. size dependency of the structure activity correlation of alkanes, Bull. Chem. Soc. Jpn. 61 (1988) 3093–3102.

[28] I. Gutman, Acyclic systems with extremal Hückel π-electron energy, Theor. Chim. Acta 45 (1977) 79–87.

[29] I. Gutman, The energy of a graph, Ber. Math.-Statist. Sekt. Forsch. Graz 22 (103) (1978) 100–105.

[30] I. Gutman, Graphs with greatest number of matching, Publ. Inst. Math. 27 (1980) 67–76.

[31] I. Gutman, The energy of a graph: old and new results, in: Algebraic combinatorics and applications (Gößweinstein, 1999), Springer, Berlin, 2001, pp. 196–211.

[32] I. Gutman, H. Deng, S. Radenković, The Estrada index: an updated survey, Zb. Rad. (Beogr.) 14(22) (Selected topics on applications of graph spectra) (2011) 155–174.

[33] I. Gutman, B. Furtula, E. O. D. Andriantiana, M. Cvetić, More trees with large energy and small size, MATCH Commun. Math. Comput. Chem. 68 (2012) 697–702.

[34] I. Gutman, B. Furtula, H. Hua, Bipartite unicyclic graphs with maximal, second-maximal, and third-maximal energy, MATCH Commun. Math. Comput. Chem. 58 (2007) 75–82.

[35] I. Gutman, B. Furtula, V. Marković, B. Glišić, Alkanes with Greatest Estrada Index, Z. Naturforsch. 62a (2007) 495–497.

[36] I. Gutman, O. E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin, 1986.

[37] I. Gutman, T. Soldatovic, A. Graovac, S. Vukovic, Approximating the total pi-electron energy by means of spectral moments, Chemical Physics Letters 334 (1) (2001) 168–172.

[38] I. Gutman, D. Vidović, Quest for molecular graphs with maximal energy: A computer experiment, J. Chem. Inf. Comput. Sci. 41 (2001) 1002–1005.

[39] I. Gutman, D. Vidovic, H. Hosoya, The relation between the eigenvalue sum and the topological index Z revisited, Bulletin of the Chemical Society of Japan 75 (8) (2002) 1723–1727.

[40] I. Gutman, T. Yamaguchi, H. Hosoya, Topological index as applied to π-electronic systems. IV. on the topological factors causing non-uniform π-electron charge distribution in non-alternant hydrocarbons, Bull. Chem. Soc. Jpn. 49 (1976) 1811–1816.

[41] I. Gutman, F. Zhang, On the ordering of graphs with respect to their matching numbers, Discrete Appl. Math. 15 (1986) 25–33.

[42] C. Heuberger, S. G. Wagner, Maximizing the number of independent subsets over trees with bounded degree, J. Graph Theory 58 (1) (2008) 49–68.

[43] C. Heuberger, S. G. Wagner, Chemical trees minimizing energy and Hosoya index, J. Math. Chem. 46 (2009) 214–230.

[44] H. Hosoya, Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons, Bull. Chem. Soc. Jpn. 44 (1971) 2332–2339.

[45] H. Hosoya, Chemical meaning of octane number analyzed by topological indices, Croat. Chem. Acta 75 (2002) 433–445.

[46] H. Hosoya, Important mathematical structures of the topological index Z for tree graphs., J. Chem. Inf. Model. 47 (3) (2007) 744–750.

[47] H. Hosoya, Mathematical meaning and importance of the topological index Z, Croat. Chem. Acta 80 (2007) 239–249.

[48] H. Hosoya, K. Hosoi, I. Gutman, A topological index for the total π-electron energy. Proof of a generalised Hückel rule for an arbitrary network, Theor. Chim. Acta 38 (1975) 37–47.

[49] H. Hosoya, K. Kawasaki, K. Mizutani, Topological index and thermodynamic properties. I. Empirical rules on the boiling point of saturated hydrocarbons, Bulletin of the Chemical Society of Japan 45 (11) (1972) 3415–3421.

[50] Y. Hou, Unicyclic graphs with minimal energy, J. Math. Chem. 29 (2001) 163–168.

[51] Y. Hou, I. Gutman, C.-W. Woo, Unicyclic graphs with maximal energy, Linear Algebra Appl. 356 (2002) 27–36, special issue on algebraic graph theory (Edinburgh, 2001).

[52] H. Hua, Bipartite unicyclic graphs with large energy, MATCH Commun. Math. Comput. Chem. 58 (2007) 57–73.

[53] B. Huo, S. Ji, X. Li, Note on unicyclic graphs with given number of pendent vertices and minimal energy, Linear Algebra and its Applications 433 (7) (2010) 1381–1387.

[54] B. Huo, S. Ji, X. Li, Y. Shi, Complete solution to a conjecture on the fourth maximal energy tree, MATCH Commun. Math. Comput. Chem. 66 (2011) 903–912.

[55] B. Huo, X. Li, Y. Shi, Complete solution to a conjecture on the maximal energy of unicyclic graphs, European Journal of Combinatorics 32 (2011) 662–673.

[56] B. Huo, X. Li, Y. Shi, Complete solution to a problem on the maximal energy of unicyclic bipartite graphs, Linear Algebra Appl. 434 (2011) 1370–1377.

[57] A. Ilić, D. Stevanović, The Estrada index of chemical trees, J. Math. Chem. 47 (1) (2010) 305–314.

[58] M. S. Jacobson, A. Kézdy, S. Seif, The poset on connected induced subgraphs of a graph need not be Sperner, Order 12 (1995) 315–318.

[59] R. Jamison, On the average number of nodes in a subtree of a tree, J. Combin. Theory Ser. B 35 (1983) 207–223.

[60] R. E. Jamison, Monotonicity of the mean order of subtrees, J. Combin. Theory Ser. B 37 (1) (1984) 70–78.

[61] M. Klazar, Twelve countings with rooted plane trees, European J. Combin. 18 (2) (1997) 195–210.

[62] A. Knopfmacher, R. F. Tichy, S. Wagner, V. Ziegler, Graphs, partitions and Fibonacci numbers, Discrete Appl. Math. 155 (2007) 1175–1187.

[63] B. Knudsen, Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree, in: Lecture Notes in Bioinformatics, vol. 2812, Springer-Verlag, 2003, pp. 433–446.

[64] N. Li, S. Li, On the extremal energies of trees, MATCH Commun. Math. Comput. Chem. 59 (2008) 291–314.

[65] S. Li, X. Li, Z. Zhu, On minimal energy and Hosoya index of unicyclic graphs, MATCH Commun. Math. Comput. Chem. 61 (2009) 325–339.

[66] S. Li, S. Wang, Enumerating the total number of subtrees of trees, preprint, `http://arxiv.org/abs/1206.2975`.

[67] S. Li, S. Wang, Further analysis on the total number of subtrees of trees, Electron. J. Combin. 19 (4) (2012) Research Paper 48, 14 pp.

[68] X. Li, Y. Shi, I. Gutman, Graph energy, Springer, New York, 2012.

[69] Y. Liu, Some results on energy of unicyclic graphs with n vertices, J. Math. Chem. 47 (2010) 1–10.

[70] J. L. Martin, M. Morin, J. D. Wagner, On distinguishing trees by their chromatic symmetric functions, J. Combin. Theory Ser. A 115 (2) (2008) 237–253.

[71] R. Merrifield, H. Simmons, The structures of molecular topological spaces, Theoretica chimica acta 55 (1980) 55–75.

[72] H. Narumi, H. Hosoya, Topological index and thermodynamic properties. ii. analysis of the topological factors on the absolute entropy of acyclic saturated hydrocarbons, Bulletin of the Chemical Society of Japan 53 (5) (1980) 1228–1237.

[73] X.-F. Pan, J.-M. Xu, C. Yang, M.-J. Zhou, Some graphs with minimum Hosoya index and maximum Merrifield-Simmons index, MATCH Commun. Math. Comput. Chem. 57 (2007) 235–242.

[74] H. Prodinger, R. F. Tichy, Fibonacci numbers of graphs, Fibonacci Quart. 20 (1982) 16–21.

[75] N. Schmuck, S. Wagner, H. Wang, Greedy trees, Caterpillars, and Wiener-type graph invariants, MATCH Commun. Math. Comput. Chem. 68 (1) (2012) 273–292.

[76] H.-Y. Shan, J.-Y. Shao, L. Zhang, C.-X. He, Proof of a conjecture on trees with large energy, MATCH Commun. Math. Comput. Chem. 68 (2012) 703–720.

[77] L. A. Székely, H. Wang, On subtrees of trees, Adv. in Appl. Math. 34 (1) (2005) 138–155.

[78] L. A. Székely, H. Wang, T. Wu, The sum of the distances between the leaves of a tree and the 'semi-regular' property, Discrete Math. 311 (13) (2011) 1197–1203.

[79] W. T. Trotter, J. L. Moore, Some theorems on graphs and posets, Discrete Math. 15 (1976) 79–84.

[80] A. Vince, H. Wang, Infinitely many trees have non-Sperner subtree poset, Order 24 (2007) 133–138.

[81] A. Vince, H. Wang, The average order of a subtree of a tree, J. Combin. Theory Ser. B 100 (2010) 161–170.

[82] S. Wagner, Extremal trees with respect to Hosoya index and Merrifield-Simmons index, MATCH Commun. Math. Comput. Chem. 57 (2007) 221–233.

[83] S. Wagner, I. Gutman, Maxima and minima of the Hosoya index and the Merrifield-Simmons index, Acta Appl. Math. 112 (2010) 323–346.

[84] H. Wang, The extremal values of the Wiener index of a tree with given degree sequence, Discrete Appl. Math. 156 (14) (2008) 2647–2654.

[85] H. Wang, Corrigendum: The extremal values of the Wiener index of a tree with given degree sequence, Discrete Appl. Math. 157 (2009) 3754.

[86] W. Yan, L. Ye, On the minimal energy of trees with a given diameter, Appl. Math. Lett. 18 (2005) 1046–1052.

[87] W. Yan, Y.-N. Yeh, Enumeration of subtrees of trees, Theoret. Comput. Sci. 369 (1-3) (2006) 256–268.

[88] A. Yu, X. Lv, The Merrifield-Simmons indices and Hosoya indices of trees with k pendant vertices, J. Math. Chem. 41 (2007) 33–43.

[89] J. Zhang, B. Zhou, J. Li, On Estrada index of trees, Linear Algebra Appl. 434 (1) (2011) 215–223.

[90] X.-D. Zhang, The Laplacian spectral radii of trees with degree sequences, Discrete Math. 308 (15) (2008) 3143–3150.

[91] X.-D. Zhang, The signless Laplacian spectral radius of graphs with given degree sequences, Discrete Appl. Math. 157 (13) (2009) 2928–2937.

[92] X.-D. Zhang, Q.-Y. Xiang, L.-Q. Xu, R.-Y. Pan, The Wiener index of trees with given degree sequences, MATCH Commun. Math. Comput. Chem. 60 (2) (2008) 623–644.

[93] X.-M. Zhang, X.-D. Zhang, Trees with given degree sequences that have minimal subtrees, preprint, `http://arxiv.org/abs/1209.0273`.

[94] X.-M. Zhang, X.-D. Zhang, D. Gray, H. Wang, The number of subtrees of trees with given degree sequence, Journal of Graph Theory, to appear.

[95] H. Zhao, X. Li, On the Fibonacci numbers of trees, Fibonacci Quart. 44 (2006) 32–38.